PCT

LD INTELLECTUAL PROPERTY ORGANIZATI
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C07K 14/79, 14/22, C12N 15/31, C12Q 1/68, A61K 39/02, 48/00

A2

(11) International Publication Number: WO 99/52947

(43) International Publication Date: 21 October 1999 (21.10.99)

(21) International Application Number: PCT/CA99/00307

(22) International Filing Date: 12 April 1999 (12.04.99)

(30) Priority Data:
09/059,584 14 April 1998 (14.04.98) US

(71) Applicant *(for all designated States except US)*: CONNAUGHT LABORATORIES LIMITED [CA/CA]; 1755 Steeles Avenue, North York, Ontario M2R 3T4 (CA).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: MYERS, Lisa, E. [CA/CA]; 187 Elizabeth Street, Guelph, Ontario N1E 2X5 (CA). SCHRYVERS, Anthony, B. [CA/CA]; 39 Edforth Road N.W., Calgary, Alberta T3A 3V8 (CA). HARKNESS, Robin, E. [CA/CA]; Apartment 1706, 640 Sheppard Avenue East, Willowdale, Ontario M2K 1B8 (CA). LOOSMORE, Sheena, M. [CA/CA]; 70 Crawford Rose Drive, Aurora, Ontario L4G 4R4 (CA). DU, Run-Pan [CA/CA]; 299 Chelwood Drive, Thornhill, Ontario L4J 7Y8 (CA). YANG, Yan-Ping [CA/CA]; Apartment 709, 120 Torresdale Avenue, Willowdale, Ontario M2R 3N7 (CA). KLEIN, Michel, H. [CA/CA]; 16 Munro Boulevard, Willowdale, Ontario M2P 1B9 (CA).

(74) Agent: STEWART, Michael, I.; Sim & McBurney, 6th floor, 330 University Avenue, Toronto, Ontario M5G 1R7 (CA).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published**
*Without international search report and to be republished upon receipt of that report.*

(54) Title: TRANSFERRIN RECEPTOR GENES OF *MORAXELLA*

(57) Abstract

Purified and isolated nucleic acid molecules are provided which encode Tbp2 proteins of *M. catarrhalis* strains M35, 3 and LES1. The nucleic acid sequence may be used to produce recombinant Tbp2 proteins of the strain of *Moraxella* free of other proteins of the *Moraxella* strain for purposes of diagnostics and medical treatment. Furthermore, the nucleic acid molecules may be used in the diagnosis of infection.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| AL | Albania | ES | Spain | LS | Lesotho | SI | Slovenia |
| AM | Armenia | FI | Finland | LT | Lithuania | SK | Slovakia |
| AT | Austria | FR | France | LU | Luxembourg | SN | Senegal |
| AU | Australia | GA | Gabon | LV | Latvia | SZ | Swaziland |
| AZ | Azerbaijan | GB | United Kingdom | MC | Monaco | TD | Chad |
| BA | Bosnia and Herzegovina | GE | Georgia | MD | Republic of Moldova | TG | Togo |
| BB | Barbados | GH | Ghana | MG | Madagascar | TJ | Tajikistan |
| BE | Belgium | GN | Guinea | MK | The former Yugoslav Republic of Macedonia | TM | Turkmenistan |
| BF | Burkina Faso | GR | Greece | | | TR | Turkey |
| BG | Bulgaria | HU | Hungary | ML | Mali | TT | Trinidad and Tobago |
| BJ | Benin | IE | Ireland | MN | Mongolia | UA | Ukraine |
| BR | Brazil | IL | Israel | MR | Mauritania | UG | Uganda |
| BY | Belarus | IS | Iceland | MW | Malawi | US | United States of America |
| CA | Canada | IT | Italy | MX | Mexico | UZ | Uzbekistan |
| CF | Central African Republic | JP | Japan | NE | Niger | VN | Viet Nam |
| CG | Congo | KE | Kenya | NL | Netherlands | YU | Yugoslavia |
| CH | Switzerland | KG | Kyrgyzstan | NO | Norway | ZW | Zimbabwe |
| CI | Côte d'Ivoire | KP | Democratic People's Republic of Korea | NZ | New Zealand | | |
| CM | Cameroon | | | PL | Poland | | |
| CN | China | KR | Republic of Korea | PT | Portugal | | |
| CU | Cuba | KZ | Kazakstan | RO | Romania | | |
| CZ | Czech Republic | LC | Saint Lucia | RU | Russian Federation | | |
| DE | Germany | LI | Liechtenstein | SD | Sudan | | |
| DK | Denmark | LK | Sri Lanka | SE | Sweden | | |
| EE | Estonia | LR | Liberia | SG | Singapore | | |

## TITLE OF INVENTION

## TRANSFERRIN RECEPTOR GENES OF *MORAXELLA*

## FIELD OF INVENTION

The present invention relates to the molecular cloning of genes encoding transferrin receptor (TfR) proteins and, in particular, to the cloning of transferrin receptor genes from *Moraxella (Branhamella) catarrhalis*.

## BACKGROUND OF THE INVENTION

*Moraxella (Branhamella) catarrhalis* bacteria are Gram-negative diplococcal pathogens which are carried asymptomatically in the healthy human respiratory tract. In recent years, *M. catarrhalis* has been recognized as an important causative agent of otitis media. In addition, *M. catarrhalis* has been associated with sinusitis, conjunctivitis, and urogenital infections, as well as with a number of inflammatory diseases of the lower respiratory tract in children and adults, including pneumonia, chronic bronchitis, tracheitis, and emphysema (refs. 1 to 8). (Throughout this application, various references are cited in parentheses to describe more fully the state of the art to which this invention pertains. Full bibliographic information for each citation is found at the end of the specification, immediately preceding the claims. The disclosures of these references are hereby incorporated by reference into the present disclosure). Occasionally, *M. catarrhalis* invades to cause septicaemia, arthritis, endocarditis, and meningitis (refs. 9 to 13).

Otitis media is one of the most common illnesses of early childhood and approximately 80% of all children suffer at least one middle ear infection before the age

of three (ref. 14). Chronic otitis media has been associated with auditory and speech impairment in children, and in some cases, has been associated with learning disabilities. Conventional treatments for otitis media include antibiotic administration and surgical procedures, including tonsillectomies, adenoidectomies, and tympanocentesis. In the United States, treatment costs for otitis media are estimated to be between one and two billion dollars per year.

In otitis media cases, *M. catarrhalis* commonly is co-isolated from middle ear fluid along with *Streptococcus pneumoniae* and non-typable *Haemophilus influenzae*, which are believed to be responsible for 50% and 30% of otitis media infections, respectively. *M. catarrhalis* is believed to be responsible for approximately 20% of otitis media infections (ref. 15). Epidemiological reports indicate that the number of cases of otitis media attributable to *M. catarrhalis* is increasing, along with the number of antibiotic-resistant isolates of *M. catarrhalis*. Thus, prior to 1970, no β-lactamase-producing *M. catarrhalis* isolates had been reported, but since the mid-seventies, an increasing number of β-lactamase-expressing isolates have been detected. Recent surveys suggest that 75% of clinical isolates produce β-lactamase (ref. 16, 26).

Iron is an essential nutrient for the growth of many bacteria. Several bacterial species, including *M. catarrhalis*, obtain iron from the host by using transferrin receptor proteins to capture transferrin. A number of bacteria including *Neisseria meningitidis* (ref. 17), *N. gonorrhoeae* (ref. 18), *Haemophilus influenzae* (ref. 19), as well as *M. catarrhalis* (ref. 20), produce outer membrane proteins which specifically bind human transferrin. The expression of these

proteins is regulated by the amount of iron in the environment.

The two transferrin receptor proteins of *M. catarrhalis*, designated transferrin binding protein 1 (Tbp1) and transferrin binding protein 2 (Tbp2), have molecular weights of 115 kDa (Tbp1) and approximately 80 to 90 kDa (Tbp2). Unlike the transferrin receptor proteins of other bacteria which have an affinity for apotransferrin, the *M. catarrhalis* Tbp2 receptors have a preferred affinity for iron-saturated (i.e., ferri-) transferrin (ref. 21).

*M. catarrhalis* infection may lead to serious disease. It would be advantageous to provide a recombinant source of transferrin binding proteins as antigens in immunogenic preparations including vaccines, carriers for other antigens and immunogens and the generation of diagnostic reagents. The genes encoding transferrin binding proteins and fragments thereof are particularly desirable and useful in the specific identification and diagnosis of *Moraxella* and for immunization against disease caused by *M. catarrhalis* and for the generation of diagnostic reagents.

There had previously been described in published PCT application WO 97/32380, assigned to Connaught Laboratories Limited, the assignee hereof, the cloning, subcloning and sequencing of nucleic acid molecules encoding transferrin receptor proteins Tbp1 and Tbp2 of certain specific strains of *Moraxella catarrhalis*, namely *M. catarrhlais* strains 4223, Q8 and R1, as well as identifying the deduced amino acid sequences of the encoded Tbp1 and Tbp2 proteins.

WO 97/32380 further describes the construction of expression plasmids for the production of recombinant Tbp1 from *M. catarrhalis* strain 4223 and of recombinant

Tbp2 from *M. catarrhalis* strains 4223 and Q8, the recombinant expression of such proteins in *E. coli*, and the extraction and purification of the expressed Tbp1 and Tbp2 proteins.

## SUMMARY OF THE INVENTION

The present invention is directed towards the provision of purified and isolated nucleic acid molecules encoding the transferrin receptor protein Tbp2 of additional strains of *Moraxella catarrhalis*, namely strains M35, 3 and LES1. As in the case of WO 97/32380, the respective genes encoding the Tbp1 and Tbp2 proteins are identified as *tbpA* and *tbpB* genes.

The nucleic acid molecules provided herein are useful for the specific detection of strains of *Moraxella* and for diagnosis of infection by *Moraxella*. The purified and isolated nucleic acid molecules provided herein, such as DNA, are also useful for expressing the *tbp* genes by recombinant DNA means for providing, in an economical manner, purified and isolated transferrin receptor proteins as well as subunits, fragments or analogs thereof.

The transferrin receptor, subunits or fragments thereof or analogs thereof, as well as nucleic acid molecules encoding the same and vectors containing such nucleic acid molecules, are useful in immunogenic compositions for vaccinating against diseases caused by *Moraxella*, the diagnosis of infection by *Moraxella* and as tools for the generation of immunological reagents.

Monoclonal antibodies or mono-specific antisera (antibodies) raised against the transferrin receptor protein, produced in accordance with aspects of the present invention, are useful for the diagnosis of infection by *Moraxella*, the specific detection of

*Moraxella* (in, for example, *in vitro* and *in vivo* assays) and for the treatment of diseases caused by *Moraxella*.

In accordance with one aspect of the present invention, there is provided a purified and isolated nucleic acid molecule encoding transferrin receptor protein Tbp2 of a strain of *Moraxella*, specifically *M. catarrhalis* strain M35, 3 or LES1.

In one preferred embodiment of the invention, the nucleic acid molecule may encode only the Tbp2 protein of the *Moraxella* strain.

The purified and isolated nucleic acid molecule preferably has a DNA sequence selected from the group consisting of (a) a DNA sequence as set out in Figure 2, 4 or 6 (SEQ ID NOS: 1, 3 or 5) or the complementary DNA sequence thereto; (b) a DNA sequence encoding an amino acid sequence as set out in Figure 2, 4 or 6 (SEQ ID NOS: 2, 4 or 6) or the complementary DNA sequence thereto.

In an additional aspect, the present invention includes a vector adapted for transformation of a host, comprising a nucleic acid molecule as provided herein. Such vector may further comprise expression means operatively coupled to the nucleic acid molecule for expression by the host of the Tbp2 protein of the respective strain of *M. catarrhalis*.

The expression means may include a promoter and a nucleic acid portion encoding a leader sequence for secretion from the host of the transferrin receptor protein. The expression means also may include a nucleic acid portion encoding a lipidation signal for expression from the host of a lipidated form of the transferrin receptor protein or the fragment. The host transformed by the expression vector may be selected from, for example, *Escherichia coli*, *Bordetella*,

*Bacillus*, *Haemophilus*, *Moraxella*, fungi, yeast or baculovirus and Semliki Forest virus expression systems may be used.

In an additional aspect of the invention, there is provided a transformed host containing an expression vector as provided herein. The invention further includes a recombinant Tbp2 protein of the specific strains of *Moraxella catarrhalis* and producible by the transformed host. Such recombinant Tbp2 proteins have a deduced amino acid sequence selected from the group consisting of those shown in Figure 2, 4 or 6 (SEQ ID NO: 2, 4 or 6).

Such recombinant transferrin receptor protein may be provided in substantially pure form according to a further aspect of the invention, which provides a method of forming a substantially pure recombinant Tbp2 protein of *Moraxella catarrhalis* strain M35, 3 or LES1, which comprises growing the transformed host provided herein to express Tbp2 protein as inclusion bodies, purifying the inclusion bodies free from cellular material and soluble proteins, solubilizing Tbp2 protein from the purified inclusion bodies, and purifying the Tbp2 protein free from other solubilized materials. The substantially pure recombinant transferrin receptor protein is generally at least about 70% pure, preferably at least about 90% pure.

In accordance with another aspect of the invention, an immunogenic composition is provided which comprises at least one active component selected from at least one nucleic acid molecule as provided herein and at least one recombinant protein as provided herein, and a pharmaceutically acceptable carrier therefor or vector therefor. The at least one active component produces an immune response when administered to a host.

The immunogenic compositions provided herein may be formulated as vaccines for *in vivo* administration to a host. For such purpose, the compositions may be formulated as a microparticle, capsule, ISCOM (immunostimulatory complex) or liposome preparation. The immunogenic composition may be provided in combination with a targeting molecule for delivery to specific cells of the immune system or to mucosal surfaces. The immunogenic compositions of the invention (including vaccines) may further comprise at least one other immunogenic or immunostimulating material and the immunostimulating material may be at least one adjuvant or at least one cytokine.

Suitable adjuvants for use in the present invention include (but are not limited to) aluminum phosphate, aluminum hydroxide, QS21, Quil A, derivatives and components thereof, ISCOM matrix, calcium phosphate, calcium hydroxide, zinc hydroxide, a glycolipid analog, an octadecyl ester of an amino acid, a muramyl dipeptide, polyphosphazene, ISCOPREP, DC-chol, DDBA and a lipoprotein.

Advantageous combinations of adjuvants are described in copending United States Patent Applications Nos. 08/261,194 filed June 16, 1994 and 08/483,856, filed June 7, 1995, assigned to the assignee hereof and the disclosures of which are incorporated herein by reference thereto (WO 95/34308).

In accordance with another aspect of the invention, there is provided a method for generating an immune response in a host, comprising the step of administering to a susceptible host, such as a human, an effective amount of the immunogenic composition provided herein. The immune response may be a humoral or a cell-mediated immune response and may provide protection against disease caused by *Moraxella*. Hosts in which protection

against disease may be conferred include primates, including humans.

In a further aspect of the invention, there is provided a live vector for delivery of Tbp2 protein to a host, comprising a vector containing the nucleic acid molecule as described above. The vector may be selected from *Salmonella*, BCG, adenovirus, poxvirus, vaccinia and poliovirus.

The nucleic acid molecules provided herein are useful in diagnostic applications. Accordingly, in a further aspect of the invention, there is provided a method of determining the presence, in a sample, of nucleic acid encoding a transferrin receptor protein of a strain of *Moraxella*, comprising the steps of:

(a) contacting the sample with a nucleic acid molecule as provided herein to produce duplexes comprising the nucleic acid molecule and any nucleic acid molecule encoding the transferrin receptor protein of a strain of *Moraxella* present in the sample and specifically hybridizable therewith; and

(b) determining the production of the duplexes.

In addition, the present invention provides a diagnostic kit for determining the presence, in a sample, of nucleic acid encoding a transferrin receptor protein of a strain of *Moraxella*, comprising:

(a) a nucleic acid molecule as provided herein;

(b) means for contacting the nucleic acid molecule with the sample to produce duplexes comprising the nucleic acid molecule and any such nucleic acid present in the sample and hybridizable with the nucleic acid molecule; and

(c) means for determining production of the duplexes.

The invention further includes the use of the nucleic acid molecules and proteins provided herein as

medicines. The invention additionally includes the use of the nucleic acid molecules and proteins provided herein in the manufacture of medicaments for protection against infection by strains of *Moraxella*.

Advantages of the present invention include:

- an isolated and purified nucleic acid molecule encoding a Tbp2 protein of specific strains of *Moraxella catarrhalis;*
- recombinantly-produced Tbp2 proteins; and
- diagnostic kits and immunological reagents for specific identification of *Moraxella*.

## BRIEF DESCRIPTION OF DRAWINGS

The present invention will be further understood from the following description with reference to the drawings, in which:

Figure 1 shows a partial restriction map of the *M. catarrhalis* strain M35 *tbpB* gene;

Figure 2 shows the nucleotide sequence of the *tbpB* gene (SEQ ID NO: 1) and deduced amino acid sequence of the Tbp2 protein of *M. catarrhalis* strain M35 (SEQ ID NO: 2);

Figure 3 shows a partial restriction map of the *tbpB* gene for *M. catarrhalis* strain 3;

Figure 4 shows the nucleotide sequence of *tbpB* gene (SEQ ID NO: 3) and the deduced amino acid sequence of the Tbp2 protein of *M. catarrhalis* strain 3 (SEQ ID NO: 4);

Figure 5 shows a partial restriction map of the *tbpB* genes for *M. catarrhalis* strain LES1;

Figure 6 shows the nucleotide sequence of the *tbpB* gene (SEQ ID NO: 5) and deduced amino acid sequence of the Tbp2 *M. catarrhalis* strain LES1 (SEQ ID NO: 6);

Figure 7 shows an alignment of the Tbp2 proteins from strains 4223 (SEQ ID NO: 7), R1 (SEQ ID NO: 8),

M35 (SEQ ID NO: 2), LES1 (SEQ ID NO: 6), Q8 (SEQ ID NO: 9) and 3 (SEQ ID NO: 4). Dots indicate identical residues and spaces have been introduced to maximize the sequence alignment. Underlining indicates those sequences conserved amongst the *M. catarrhalis* Tbp2 proteins and those from *A. pleuropneumoniae*, *H. influenzae*, *N. gonorrhoeae*, *N. meningitidis* and *P. haemolytica* (SEQ ID NOS: 7, 8 and 9 are disclosed in WO 97/32380);

Figure 8 shows the nucleotide and deduced amino acid sequences of the *M. catarrhalis* strain 4223 *tbpA* - *orf3* - *tbpB* gene locus (SEQ ID NO: 10 - entire gene locus; SEQ ID NO: 11 - *tbpA* coding sequence; SEQ ID NO: 12 - deduced amino acid sequence of TbpA; SEQ ID NO: 13 - *orf3* coding sequence; SEQ ID NO: 14 - deduced amino acid sequence of ORF3; SEQ ID NO: 15 - *tbpB* coding sequence; SEQ ID NO: 7 - deduced amino acid sequence of Tbp2);

Figure 9 shows an alignment of the ORF3 proteins from *M. catarrhalis* strains 4223 (SEQ ID NO: 14) and Q8 (SEQ ID NO: 16). Dots indicate identical residues;

Figure 10 shows a restriction map of clone LEM3-24 the construction of which is described in WO 97/32380 (ATCC deposit No. 97,381 deposited December 4, 1995) showing the location of the *orf3* gene in addition to the *tbpA* and *tbpB* genes of *M. catarrhalis* strain 4223 (cf. Figure 2 of WO 96/32380); and

Figure 11 shows a restriction map of clone SLRD-A the construction of which is described in WO 97/32380 (ATCC deposit No. 97,381 deposited December 4, 1995), showing the locations of the *orf3* gene in addition to the *tbpA* and *tbpB* genes of *M. catarrhalis* strain Q8 (cf. Figure 7 of WO 97/32380).

## GENERAL DESCRIPTION OF THE INVENTION

*Moraxella catarrhalis* strains M35, 3 and LES1 may be conveniently used to provide the purified and isolated nucleic acid, which may be in the form of DNA molecules, comprising at least a portion of the nucleic acid coding for a Tbp2 protein of the strain. Strains 4223, LES1 and M35 are all derived from patients with otitis media while strains 3, R1 and Q8 were from sputum or bronchial secretions.

The *tbpB* genes from *M. catarrhalis* M35, 3 and LES1 were cloned and sequenced herein, following generally the procedures described in WO 97/32380. Strain 3 is a clinical isolate provided by Dr. T. Murphy (State University of New York, Buffalo, New York); strain M35 was obtained from Dr. G.D. Campbell (Louisiana State University, Shreveport, Louisiana) and strain LES1 was obtained from Dr. L. Stanfors (University of Tromso, Finland).

Figures 2, 4 and 6 show the nucleotide sequences of the respective *tbpB* genes (SEQ ID NO: 1, 3 or 5) and deduced amino acid sequence of the Tbp2 protein (SEQ ID NO: 2, 4 or 6) of the *M. catarrhalis* strains M35, 3 and LES1, respectively. Regions of homology are evident between the *M. catarrhalis* Tbp2 amino acid sequences determined herein and those previously determined in WO 97/32380, as shown in the comparative alignment of Figure 7 (SEQ ID NOS: 7, 8, 2, 6, 9 and 4) and between the *M. catarrhalis* Tbp2 amino acid sequences. Underlining in Figure 7 indicates those sequences which are conserved among the *M. catarrhalis* Tbp2 proteins and those of *A. pleuropneumoniae*, *H. influenzae*, *N. gonorrhoeae*, *N. meningitidis* and *P. haemolytica*.

Sequence analysis of the nucleotide acid and amino acid sequences of the Tbp2 proteins described herein

and in WO 97/32380 indicated that at least two families could be identified for *M. catarrhalis tbpB* genes, one comprising strains 4223, R1 and M35 and other comprising strains Q8 and 3, with strain LES1 being equally related to both families. Anti-rTbp2 bactericidal antibody activity (Table 1) correlated with the putative gene families identified by sequencing.

Additional sequence analysis of the entire *M. catarrhalis* strains 4223 and Q8 *tbpA* - *tbpB* locus gene sequence (Figure 8) identified an intergenic open reading frame termed "*orf3*" (SEQ ID NO: 13, SEQ ID NO: 14, ORF3 amino acid sequence), (see also Figures 10 and 11 for location of orf3). The encoded ORF3 proteins from 4223 and Q8 are 98% identical, as seen from the sequence alignment of Figure 9 (SEQ ID NOS: 14, 16).

Cloned *tbpB* genes may be expressed in *E. coli* to produce recombinant Tbp2 proteins free of other *Moraxella* proteins. These recombinant proteins may be purified and used for immunization.

The Tbp2 proteins provided herein are useful as a diagnostic reagent, as an antigen for the generation of anti-transferrin protein binding antibodies, as an antigen for vaccination against the disease caused by species of *Moraxella* and for detecting infection by *Moraxella* and other such bacteria.

The Tbp2 proteins provided herein may also be used as a carrier protein for haptens, polysaccharides or peptides to make conjugate vaccines against antigenic determinants unrelated to transferrin binding proteins. In additional embodiments of the present invention, therefore, the Tbp2 proteins as provided herein may be used as a carrier molecule to prepare chimeric molecules and conjugate vaccines (including glycoconjugates)

against pathogenic bacteria, including encapsulated bacteria. Thus, for example, glycoconjugates of the present invention may be used to confer protection against disease and infection caused by any bacteria having polysaccharide antigens including lipooligosaccharides (LOS) and PRP. Such bacterial pathogens may include, for example, *Haemophilus influenzae, Streptococcus pneumoniae, Escherichia coli, Neisseria meningitidis, Salmonella typhi, Streptococcus mutans, Cryptococcus neoformans, Klebsiella, Staphylococcus aureus* and *Pseudomonas aeruginosa*. Particular antigens which can be conjugated to Tbp2 proteins and methods to achieve such conjugations are described in U.S. Patent Application No. 08/433,522 filed November 23, 1993 (WO 94/12641), assigned to the assignee hereof and the disclosure of which is hereby incorporated by reference thereto.

In another embodiment, the carrier function of the Tbp2 proteins may be used, for example, to induce an immune response against abnormal polysaccharides of tumour cells, or to produce anti-tumour antibodies that can be conjugated to chemotherapeutic or bioactive agents.

The invention extends to transferrin binding proteins from *Moraxella catarrhalis* for use as an active ingredient in a vaccine against disease caused by infection with *Moraxella*. The invention also extends to a pharmaceutical vaccinal composition containing transferrin binding proteins from *Moraxella catarrhalis* and optionally, a pharmaceutically acceptable carrier and/or diluent.

In a further aspect the invention provides the use of transferrin binding proteins for the preparation of a

pharmaceutical vaccinal composition for immunization against disease caused by infection with *Moraxella*.

It is clearly apparent to one skilled in the art, that the various embodiments of the present invention have many applications in the fields of vaccination, diagnosis, treatment of, for example, *Moraxella* infections and the generation of immunological and other diagnostic reagents. A further non-limiting discussion of such uses is further presented below.

**1. Vaccine Preparation and Use**

Immunogenic compositions, suitable to be used as vaccines, may be prepared from immunogenic transferrin receptor proteins, analogs and fragments thereof encoded by the nucleic acid molecules as well as the nucleic acid molecules disclosed herein. The vaccine elicits an immune response which produces antibodies, including anti-transferrin receptor antibodies and antibodies that are opsonizing or bactericidal. Should the vaccinated subject be challenged by *Moraxella*, the antibodies bind to the transferrin receptor and thereby prevent access of the bacteria to an iron source which is required for viability. Furthermore, opsonizing or bactericidal anti-transferrin receptor antibodies may also provide protection by alternative mechanisms.

Immunogenic compositions, including vaccines, may be prepared as injectables, as liquid solutions or emulsions. The transferrin receptor proteins, analogs and fragments thereof and encoding nucleic acid molecules may be mixed with pharmaceutically acceptable excipients which are compatible with the transferrin receptor proteins, fragments, analogs or nucleic acid molecules. Such excipients may include water, saline, dextrose, glycerol, ethanol, and combinations thereof. The immunogenic compositions and vaccines may further

contain auxiliary substances, such as wetting or emulsifying agents, pH buffering agents, or adjuvants, to enhance the effectiveness of the vaccines. Immunogenic compositions and vaccines may be administered parenterally, by injection subcutaneously, intradermally or intramuscularly. Alternatively, the immunogenic compositions provided according to the present invention, may be formulated and delivered in a manner to evoke an immune response at mucosal surfaces. Thus, the immunogenic composition may be administered to mucosal surfaces by, for example, the nasal or oral (intragastric) routes. The immunogenic composition may be provided in combination with a targeting molecule for delivery to specific cells of the immune system or to mucosal surfaces. Some such targeting molecules include vitamin B12 and fragments of bacterial toxins, as described in WO 92/17167 (Biotech Australia Pty. Ltd.), and monoclonal antibodies, as described in U.S. Patent No. 5,194,254 (Barber et al). Alternatively, other modes of administration, including suppositories and oral formulations, may be desirable. For suppositories, binders and carriers may include, for example, polyalkalene glycols or triglycerides. Oral formulations may include normally employed incipients such as, for example, pharmaceutical grades of saccharine, cellulose and magnesium carbonate. These compositions may take the form of solutions, suspensions, tablets, pills, capsules, sustained release formulations or powders and contain about 1 to 95% of the transferrin receptor proteins, fragments, analogs and/or nucleic acid molecules.

The vaccines are administered in a manner compatible with the dosage formulation, and in such amount as will be therapeutically effective, protective and immunogenic. The quantity to be administered

depends on the subject to be treated, including, for example, the capacity of the individual's immune system to synthesize antibodies, and, if needed, to produce a cell-mediated immune response. Precise amounts of active ingredient required to be administered depend on the judgment of the practitioner. However, suitable dosage ranges are readily determinable by one skilled in the art and may be of the order of micrograms of the transferrin receptor proteins, analogs and fragments thereof and/or nucleic acid molecules. Suitable regimes for initial administration and booster doses are also variable, but may include an initial administration followed by subsequent administrations. The dosage of the vaccine may also depend on the route of administration and will vary according to the size of the host.

The nucleic acid molecules encoding the transferrin receptor of *Moraxella* may be used directly for immunization by administration of the DNA directly, for example, by injection for genetic immunization or by constructing a live vector, such as *Salmonella*, BCG, adenovirus, poxvirus, vaccinia or poliovirus containing the nucleic acid molecules. A discussion of some live vectors that have been used to carry heterologous antigens to the immune system is contained in, for example, O'Hagan (ref. 22). Processes for the direct injection of DNA into test subjects for genetic immunization are described in, for example, Ulmer et al. (ref. 23).

Immunogenicity can be significantly improved if the antigens are co-administered with adjuvants, commonly used as an 0.05 to 1.0 percent solution in phosphate - buffered saline. Adjuvants enhance the immunogenicity of an antigen but are not necessarily immunogenic themselves. Adjuvants may act by retaining the antigen

locally near the site of administration to produce a depot effect facilitating a slow, sustained release of antigen to cells of the immune system. Adjuvants can also attract cells of the immune system to an antigen depot and stimulate such cells to elicit immune responses.

Immunostimulatory agents or adjuvants have been used for many years to improve the host immune responses to, for example, vaccines. Intrinsic adjuvants, such as lipopolysaccharides, normally are the components of killed or attenuated bacteria used as vaccines. Extrinsic adjuvants are immunomodulators which are typically non-covalently linked to antigens and are formulated to enhance the host immune responses. Thus, adjuvants have been identified that enhance the immune response to antigens delivered parenterally. Some of these adjuvants are toxic, however, and can cause undesirable side-effects, making them unsuitable for use in humans and many animals. Indeed, only aluminum hydroxide and aluminum phosphate (collectively commonly referred to as alum) are routinely used as adjuvants in human and veterinary vaccines. The efficacy of alum in increasing antibody responses to diphtheria and tetanus toxoids is well established and an HBsAg vaccine has been adjuvanted with alum. While the usefulness of alum is well established for some applications, it has limitations. For example, alum is ineffective for influenza vaccination and inconsistently elicits a cell mediated immune response. The antibodies elicited by alum-adjuvanted antigens are mainly of the IgG1 isotype in the mouse, which may not be optimal for protection by some vaccinal agents.

A wide range of extrinsic adjuvants can provoke potent immune responses to antigens. These include saponins complexed to membrane protein antigens (immune

stimulating complexes), pluronic polymers with mineral oil, killed mycobacteria and mineral oil, Freund's complete adjuvant, bacterial products, such as muramyl dipeptide (MDP) and lipopolysaccharide (LPS), as well as lipid A, and liposomes.

To efficiently induce humoral immune responses (HIR) and cell-mediated immunity (CMI), immunogens are often emulsified in adjuvants. Many adjuvants are toxic, inducing granulomas, acute and chronic inflammations (Freund's complete adjuvant, FCA), cytolysis (saponins and pluronic polymers) and pyrogenicity, arthritis and anterior uveitis (LPS and MDP). Although FCA is an excellent adjuvant and widely used in research, it is not licensed for use in human or veterinary vaccines because of its toxicity.

Desirable characteristics of ideal adjuvants include:

(1) lack of toxicity;

(2) ability to stimulate a long-lasting immune response;

(3) simplicity of manufacture and stability in long-term storage;

(4) ability to elicit both CMI and HIR to antigens administered by various routes, if required;

(5) synergy with other adjuvants;

(6) capability of selectively interacting with populations of antigen presenting cells (APC);

(7) ability to specifically elicit appropriate $T_H1$ or $T_H2$ cell-specific immune responses; and

(8) ability to selectively increase appropriate antibody isotype levels (for example, IgA) against antigens.

U.S. Patent No. 4,855,283 granted to Lockhoff et al on August 8, 1989, which is incorporated herein by reference thereto, teaches glycolipid analogues

including N-glycosylamides, N-glycosylureas and N-glycosylcarbamates, each of which is substituted in the sugar residue by an amino acid, as immuno-modulators or adjuvants. Thus, Lockhoff et al. 1991 (ref. 24) reported that N-glycolipid analogs displaying structural similarities to the naturally-occurring glycolipids, such as glycophospholipids and glycoglycerolipids, are capable of eliciting strong immune responses in both herpes simplex virus vaccine and pseudorabies virus vaccine. Some glycolipids have been synthesized from long chain-alkylamines and fatty acids that are linked directly with the sugars through the anomeric carbon atom, to mimic the functions of the naturally occurring lipid residues.

U.S. Patent No. 4,258,029 granted to Moloney, assigned to the assignee hereof and incorporated herein by reference thereto, teaches that octadecyl tyrosine hydrochloride (OTH) functions as an adjuvant when complexed with tetanus toxoid and formalin inactivated type I, II and III poliomyelitis virus vaccine. Also, Nixon-George et al. 1990, (ref. 25) reported that octadecyl esters of aromatic amino acids complexed with a recombinant hepatitis B surface antigen, enhanced the host immune responses against hepatitis B virus.

**2. Immunoassays**

The transferrin receptor proteins, analogs and/or fragments thereof of the present invention are useful as immunogens, as antigens in immunoassays including enzyme-linked immunosorbent assays (ELISA), RIAs and other non-enzyme linked antibody binding assays or procedures known in the art for the detection of anti-*Moraxella*, transferrin receptor protein antibodies. In ELISA assays, the transferrin receptor protein, analogs and/or fragments corresponding to portions of TfR protein, are immobilized onto a selected surface, for

example, a surface capable of binding proteins or peptides such as the wells of a polystyrene microtiter plate. After washing to remove incompletely adsorbed transferrin receptor, analogs and/or fragments, a non-specific protein such as a solution of bovine serum albumin (BSA) or casein that is known to be antigenically neutral with regard to the test sample may be bound to the selected surface. This allows for blocking of nonspecific adsorption sites on the immobilizing surface and thus reduces the background caused by non-specific bindings of antisera onto the surface.

The immobilizing surface is then contacted with a sample, such as clinical or biological materials, to be tested in a manner conducive to immune complex (antigen/antibody) formation. This procedure may include diluting the sample with diluents, such as BSA, bovine gamma globulin (BGG) and/or phosphate buffered saline (PBS)/Tween. The sample is then allowed to incubate for from about 2 to 4 hours, at temperatures such as of the order of about 25° to 37°C. Following incubation, the sample-contacted surface is washed to remove non-immunocomplexed material. The washing procedure may include washing with a solution such as PBS/Tween or a borate buffer.

Following formation of specific immunocomplexes between the test sample and the bound transferrin receptor protein, analogs and/or fragments and subsequent washing, the occurrence, and even amount, of immunocomplex formation may be determined by subjecting the immunocomplex to a second antibody having specificity for the first antibody. If the test sample is of human origin, the second antibody is an antibody having specificity for human immunoglobulins and in general IgG. To provide detecting means, the second

antibody may have an associated activity such as an enzymatic activity that will generate, for example, a color development upon incubating with an appropriate chromogenic substrate. Quantification may then achieved by measuring the degree of color generation using, for example, a spectrophotometer.

**3. Use of Sequences as Hybridization Probes**

The nucleotide sequences of the present invention, comprising the sequence of the transferrin receptor gene, now allow for the identification and cloning of the transferrin receptor genes from any species of *Moraxella*.

The nucleotide sequences comprising the sequence of the transferrin receptor genes of the present invention are useful for their ability to selectively form duplex molecules with complementary stretches of other TfR genes. Depending on the application, a variety of hybridization conditions may be employed to achieve varying degrees of selectivity of the probe toward the other TfR genes. For a high degree of selectivity, relatively stringent conditions are used to form the duplexes, such as low salt and/or high temperature conditions, such as provided by 0.02 M to 0.15 M NaCl at temperatures of between about 50°C to 70°C. For some applications, less stringent hybridization conditions are required such as 0.15 M to 0.9 M salt, at temperatures ranging from between about 20°C to 55°C. Hybridization conditions can also be rendered more stringent by the addition of increasing amounts of formamide, to destabilize the hybrid duplex. Thus, particular hybridization conditions can be readily manipulated, and will generally be a method of choice depending on the desired results. In general, convenient hybridization temperatures in the presence of 50%

formamide are: 42°C for a probe which is 95 to 100% homologous to the target fragment, 37°C for 90 to 95% homology and 32°C for 85 to 90% homology.

In a clinical diagnostic embodiment, the nucleic acid sequences of the TfR genes of the present invention may be used in combination with an appropriate means~~,~~ such as a label, for determining hybridization. A wide variety of appropriate indicator means are known in the art, including radioactive, enzymatic or other ligands, such as avidin/biotin and digoxigenin-labelling, which are capable of providing a detectable signal. In some diagnostic embodiments, an enzyme tag such as urease, alkaline phosphatase or peroxidase, instead of a radioactive tag may be used. In the case of enzyme tags, colorimetric indicator substrates are known which can be employed to provide a means visible to the human eye or spectrophotometrically, to identify specific hybridization with samples containing TfR gene sequences.

The nucleic acid sequences of TfR genes of the present invention are useful as hybridization probes in solution hybridizations and in embodiments employing solid-phase procedures. In embodiments involving solid-phase procedures, the test DNA (or RNA) from samples, such as clinical samples, including exudates, body fluids (e. g., serum, amniotic fluid, middle ear effusion, sputum, bronchoalveolar lavage fluid) or even tissues, is adsorbed or otherwise affixed to a selected matrix or surface. The fixed, single-stranded nucleic acid is then subjected to specific hybridization with selected probes comprising the nucleic acid sequences of the TfR genes or fragments thereof of the present invention under desired conditions. The selected conditions will depend on the particular circumstances

based on the particular criteria required depending on, for example, the G+C contents, type of target nucleic acid, source of nucleic acid, size of hybridization probe etc. Following washing of the hybridization surface so as to remove non-specifically bound probe molecules, specific hybridization is detected, or even quantified, by means of the label. It is preferred to select nucleic acid sequence portions which are conserved among species of *Moraxella*. The selected probe may be at least 18bp and may be in the range of about 30 to 90 bp.

**4. Expression of the Transferrin Receptor Genes**

Plasmid vectors containing replicon and control sequences which are derived from species compatible with the host cell may be used for the expression of the transferrin receptor genes in expression systems. The vector ordinarily carries a replication site, as well as marking sequences which are capable of providing phenotypic selection in transformed cells. For example, *E. coli* may be transformed using pBR322 which contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR322 plasmid, or other microbial plasmid or phage, must also contain, or be modified to contain, promoters which can be used by the host cell for expression of its own proteins.

In addition, phage vectors containing replicon and control sequences that are compatible with the host can be used as a transforming vector in connection with these hosts. For example, the phage in lambda GEM™-11 may be utilized in making recombinant phage vectors which can be used to transform host cells, such as *E. coli* LE392.

Promoters commonly used in recombinant DNA construction include the β-lactamase (penicillinase) and lactose promoter systems and other microbial promoters, such as the T7 promoter system as described in U.S. Patent No. 4,952,496. Details concerning the nucleotide sequences of promoters are known, enabling a skilled worker to ligate them functionally with genes. The particular promoter used will generally be a matter of choice depending upon the desired results. Hosts that are appropriate for expression of the transferrin receptor genes, fragments, analogs or variants thereof, may include *E. coli, Bacillus* species, *Haemophilus*, fungi, yeast, *Moraxella*, *Bordetella*, or the baculovirus expression system may be used.

In accordance with this invention, it is preferred to make the transferrin receptor protein, fragment or analog thereof, by recombinant methods, particularly since the naturally occurring TfR protein as purified from a culture of a species of *Moraxella* may include trace amounts of toxic materials or other contaminants. This problem can be avoided by using recombinantly produced TfR protein in heterologous systems which can be isolated from the host in a manner to minimize contaminants in the purified material. Particularly desirable hosts for expression in this regard include Gram positive bacteria which do not have LPS and are, therefore, endotoxin free. Such hosts include species of *Bacillus* and may be particularly useful for the production of non-pyrogenic transferrin receptor, fragments or analogs thereof.

**EXAMPLES**

The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific

Examples. These Examples are described solely for purposes of illustration and are not intended to limit the scope of the invention. Changes in form and substitution of equivalents are contemplated as circumstances may suggest or render expedient. Although specific terms have been employed herein, such terms are intended in a descriptive sense and not for purposes of limitations.

Methods of molecular genetics, protein biochemistry and immunology used but not explicitly described in this disclosure and these Examples are amply reported in the scientific literature and are well within the ability of those skilled in the art.

**Example 1**

This Example illustrates the preparation of chromosomal DNA from *M. catarrhalis* strain M35, following the procedure described in WO 97/32380 for strains 4223 and Q8 (Example 2).

*M. catarrhalis* isolate M35 was inoculated into 100 ml of BHI broth, and incubated for 18 hr at 37°C with shaking. The cells were harvested by centrifugation at 10,000 x g for 20 min. The pellet was used for extraction of *M. catarrhalis* M35 chromosomal DNA.

The cell pellet was resuspended in 20 ml of 10 mM Tris-HCl (pH 7.5)-1.0 mM EDTA (TE). Pronase and SDS were added to final concentrations of 500 μg/ml and 1.0%, respectively, and the suspension was incubated at 37°C for 2 hr. After several sequential extractions with phenol, phenol:chloroform (1:1), and chloroform:isoamyl alcohol (24:1), the aqueous extract was dialysed, at 4°C, against 1.0 M NaCl for 4 hr, and against TE (pH 7.5) for a further 48 hr with three buffer changes. Two volumes of ethanol were added to the dialysate, and the DNA was spooled onto a glass rod.

The DNA was allowed to air-dry, and was dissolved in 3.0 ml of water. Concentration was estimated, by UV spectrophotometry, to be about 290 μg/ml. This procedure was repeated for the preparation of chromosomal DNA from *M. catarrhalis* strain 3 and LES1.

**Example 2**

This Example illustrates the construction of a *M. catarrhalis* strain M35 chromosomal library in EMBL3.

A series of *Sau*3A restriction digests of chromosomal DNA from *M. catarrhalis* M35, prepared as described in Example 1, in final volumes of 10 μL each, were carried out in order to optimize the conditions necessary to generate maximal amounts of restriction fragments within a 15 to 23 kb size range. Using the optimized digestion conditions, a large-scale digestion was set up in a 100 μL volume, containing the following: 50 μL of chromosomal DNA (290 μg/ml), 33 μL water, 10 μL 10X *Sau*3A buffer (New England Biolabs), 1.0 μL BSA (10 mg/ml, New England Biolabs), and 6.3 μL *Sau*3A (0.04 U/μL). Following a 15 min. incubation at 37°C, the digestion was terminated by the addition of 10 μL of 100 mM Tris-HCl (pH 8.0)-10 mM EDTA-0.1% bromophenol blue-50% glycerol (loading buffer). Digested DNA was electrophoresed through a 0.5% agarose gel in 40 mM Tris acetate-2 mM $Na_2EDTA.2H_2O$ (pH8.5)(TAE buffer) at 50 V for 6 hr. The region containing restriction fragments within a 15 to 23 kb molecular size range was excised from the gel, and placed into dialysis tubing containing 3.0 ml of TAE buffer. DNA was electroeluted from the gel fragment by applying a field strength of 1.0 V/cm for 18 hr. Electroeluted DNA was extracted once each with phenol and phenol:chloroform (1:1), and

precipitated with ethanol. The dried DNA was dissolved in 5.0 μL water.

Size-fractionated chromosomal DNA was ligated with *Bam*HI-digested EMBL3 arms (Promega), using T4 DNA ligase in a final volume of 9 μL. The entire ligation mixture was packaged into lambda phage using a commercial packaging kit (Amersham), following manufacturer's instructions.

The packaged DNA library was amplified on solid media. 0.1 ml aliquots of *Escherichia coli* strain NM539 in 10 mM $MgSO_4$ ($OD_{260}$ = 0.5) were incubated at 37°C for 15 min. with 15 to 25 μL of the packaged DNA library. Samples were mixed with 3 ml of 0.6% agarose containing 1.0% BBL trypticase peptone-0.5% NaCl (BBL top agarose), and mixtures were plated onto 1.5% agar plates containing 1.0% BBL trypticase peptone-0.5% NaCl, and incubated at 37°C for 18 hr. 3 ml quantities of 50 mM Tris-HCl (pH 7.5)-8 mM magnesium sulfate heptahydrate-100 mM NaCl-0.01% (w/v) gelatin (SM buffer) were added to each plate, and plates were left at 4°C for 7 hr. SM buffer containing phage was collected from the plates, pooled together, and stored in a screwcap tube at 4°C, with chloroform.

**Example 3**

This Example illustrates screening of the *M. catarrhalis* strain M35 library.

The EMBL3/M35 library, prepared as described in Example 2, was plated onto LE392 cells on YT plates using 0.7% top agar in YT as overlay. Plaques were lifted onto nitrocellulose filters and the filters were probed with oligonucleotide probes labelled with $^{32}P\alpha$-dCTP (Random Primed DNA labeling kit, Boehringer Mannheim). The pre-hybridization was performed in sodium chloride/sodium citrate (SSC) buffer (ref. 27) at

37°C for 1 hour and the hybridization was performed at 42°C overnight. The probes were based upon an internal sequence of 4223 *tbpA*:

I R D L T R Y D P G

(SEQ ID No. 17)

4236-RD 5' ATTCGAGACTTAACACGCTATGACCCTGGC 3'

(Seq ID No 18)

4237-RD 5' ATTCGTGATTTAACTCGCTATGACCCTGGT 3'

(Seq ID No 19).

Putative plaques were re-plated and submitted to second and third rounds of screening using the same procedures.

Phage clone M35-2.3 was found to contain a 13 kb insert of the M35 *tfr* genes. The *tbpB* gene was localized to a 7.5 kb NheI - Sal I fragment by restriction enzyme and Southern blot analyses and was subcloned into pBR328 for sequence analysis, generating plasmid pLEM40.

A partial restriction map of the M35 *tbpB* gene is shown in Figure 1. The nucleotide and deduced amino acid sequences of the M35 *tbpB* gene are shown in Figure 2. The M35 *tbpB* gene encodes a 706 amino acid protein of molecular weight 76.5 kDa. When the M35 TbpB sequence was aligned with the 4223 TbpB protein (Figure 7), it was found to be 86% identical and 90% similar.

**Example 4**

This Example illustrates the PCR amplification of the *tbpB* genes from *M. catarrhalis* strains 3 and LES1, following the procedure described in WO 97/32380 for *M. catarrhalis* strain R1.

Oligonucleotide primers were based upon the following sequences, which are found in the intergenic regions surrounding *M. catarrhalis* strain 4223 *tbpB*:

5′ GATGGGATAAGCACGCCCTACTT 3′ (SEQ ID NO: 20)

sense primer (4940)

5' CCCATCAGCCAAACAAACATTGTGT 3' (SEQ ID NO: 21)
antisense primer (4967)

PCR amplification was performed in buffer containing 100 mM Tris-HCI (pH 8.9), 25 mM KCI, 5 mM $(NH_4)_2SO_4$ and 2 mM $MgSO_4$. Each 100 µl reaction mixture contained 10 ng of chromosomal DNA from strains 3 and LES1, prepared following the procedure of Example 1, 1 µg each primer, 2.5 U Pwo DNA polymerase (Boehringer Mannheim) and 0.2 mM dNTPs (Perkin Elmer, Foster City, California). The cycling conditions were 25 cycles of 95°C for 30 sec, 45°C for 1.0 min and 72°C for 2.0 min, followed by a 10 min elongation at 72°C. Specific 2.4 kb fragments were amplified and DNA was purified for direct sequencing by agarose gel extraction, using a Geneclean kit (Bio 101 Inc., Vista, California). Plasmid DNA for sequencing was prepared using a Qiagen Plasmid Midi kit (Qiagen, Chatsworth, California). DNA samples were sequenced using an ABI model 373A DNA sequencer using dye terminator chemistry. Oligonucleotide primers of 17 to 25 bases in length were used to sequence both strands of the genes.

Partial restriction maps of the *M. catarrhalis* strains 3 and LES1 *tbpB* genes are shown in Figures 3 and 5 respectively. The nucleotide and deduced amino acid sequences of the strain 3 and LES1 *tbpB* genes are shown in Figures 4 and 6, respectively. The strain 3 *tbpB* gene encodes a 712 amino acid protein of molecular weight 76.9 kDa, which is more closely related to the strain Q8 Tbp2 protein than to the 4223 Tbp2 protein (Figure 7). The Q8 and strain 3 Tbp2 proteins are 71% identical and 79% similar, whereas the 4223 and strain 3 Tbp2 proteins are 51% identical and 64% similar. The strain LES1 *tbpB* gene encodes a 713 amino acid protein

of molecular weight 76.8 kDa which is 63% identical to both the 4223 and Q8 Tbp2 proteins.

From the sequence analysis presented herein and in further consideration of the sequences presented in WO 98/32380, there appear to be at least two gene families which can be identified for *M. catarrhalis tbpB*, one comprising strains 4223, R1 and M35 and the other comprising strains Q8 and 3, with strain LES1 being equally related to both families. This novel finding is similar to that of the *N. meningitidis tbpB* genes which can be divided into two sub-groups (ref. 28). There is limited sequence homology among the amino acid sequences of the *M. catarrhalis* Tbp2 proteins previously identified in WO 98/32380 and in this application and those from other organisms, such as *Actinobacillus pleuropneumoniae*, *H. influenzae*, *N. gonorrhoeae*, *N. meningitidis* and *P. haemolytical* (ref. 29). The homology is scattered in small peptide motifs throughout the sequence and is illustrated by underlining in Figure 7. The conserved LEGGFYG (SEQ ID NO: 22) epitope was present, as found in Tbp2 for other *M. catarrhalis* strains as well as the *H. influenzae* and *N. meningitidis* Tbp2 proteins.

**Example 5**

This Example illustrates the bactericidal antibody activity of guinea pig anti-4223 rTbp2 and anti-Q8 rTbp2 antibodies, prepared as described in WO 97/32380 (Example 14), and confirmation of the gene families of *tbpB* genes.

The bactericidal antibody assay was performed as described by Yang et al. (ref. 30). Briefly, several *M. catarrhalis* strains were grown to an $OD_{578}$ of 0.5 in BHI medium containing 25 mM EDDA. The bacteria were diluted so that the pre-bleed control plates contained

100 to 300 cfu. Guinea pig anti-rTbp2 antisera and pre-bleed controls, prepared as described in Example 14 of WO 97/32380, were heated to 56°C for 30 min to inactivate endogenous complement and were diluted 1:64 with veronal buffer containing 0.1% BSA (VBS). Guinea pig complement was diluted 1:10 in VBS. Twenty-five µl each of diluted antiserum, bacteria and complement were added to duplicate wells of a 96 well microtiter plate. The plates were incubated at 37°C for 60 min, gently shaking at 70 rpm on a rotary platform. Fifty µl of each reaction mixture were plated onto Mueller Hinton agar plates which were incubated at 37°C for 24 h, then room temperature for 24 h, before the bacteria were counted. Antisera were determined to be bactericidial if ≥50% of bacteria were killed compared with negative controls. Each assay was repeated at least twice in duplicate. The assay was performed using both the anti-Tbp2 antisera from both 4223 and Q8 strains against a number of different strains of *Moraxella catarrhalis*. The strains tested are identified and the results obtained are shown in Table 1.

The anti-rTbp2 bactericidal antibody activity shown in Table 1 corelates with the putative gene families identified by sequencing, as described in Example 4. Anti-4223 rTbp2 antibody kills those strains within its own family, i.e. 4223, R1 and M35, while anti-Q8 rTbp2 antibody kills those strains within its family, i.e. Q8, 3 and LES1. The anti-4223 rTbp2 antibody also killed strains VH-9, H-04 and ATCC 25240 indicating that the latter strains may be part of the 4223 family. Strain H-04 was also killed by anti-Q8 rTbp2 antibody.

**Example 6**

This Example illustrates the sequence analysis of the open reading frame (ORF) within the intergenic region between *M. catarrhalis tbpA* and *tbpB*.

The intergenic region was sequenced for strains 4223 and Q8 and a single open reading frame was identified. This *orf*, identified as *orf3*, was located about 1 kb downstream of *tbpA* and about 273 bp upstream of *tbpB* in each genome (Figure 10 - strain 4223; Figure 11 - strain Q8). The nucleotide and deduced amino acid sequences of the entire 4223 *tbpA - orf3 - tbpB* gene loci are shown in Figure 8. The encoded 4223 and Q8 ORF3 proteins are 98% identical, 512 amino acid proteins, of molecular weight 58.1 kDa and 57.9 kDa, respectively. The alignment of the ORF3 protein sequences is shown in Figure 9.

**SUMMARY OF THE DISCLOSURE**

In summary of this disclosure, the present invention provides purified and isolated DNA molecules containing transferrin receptor genes of specific strains of *Moraxella catarrhalis*, the sequences of these transferrin receptor genes, and the derived amino acid sequences of the Tbp2 proteins encoded thereby. The genes and DNA sequences are useful for diagnosis, immunization, and the generation of diagnostic and immunological reagents. Immunogenic compositions, including vaccines, based upon expressed recombinant Tbp1 and/or Tbp2, portions thereof, or analogs thereof, can be prepared for prevention of diseases caused by *Moraxella*. Modifications are possible within the scope of this invention.

## TABLE I

Bactericidal antibody activity of guinea pig anti-rTbpB antisera

| *M. catarrhalis* strain | Bactericidal Antibody Activity* | |
|---|---|---|
| | Anti-4223 rTbp2 | Anti-Q8 rTbp2 |
| 4223 | ++ | - |
| M35 | ++ | - |
| R1 | ++ | - |
| LES1 | - | + |
| Q8 | - | ++ |
| 3 | - | ± |
| VH-9 | ++ | - |
| H-04 | ++ | ++ |
| ATCC 25240 | ** | - |

* killing by antiserum diluted 1:64 compared to negative controls: - indicates 0 to 25% killing; ± indicates 26 to 49%; + indicates 50 to 75%; ++ indicates 76 to 100% killing.

## REFERENCES

1. Brorson, J-E., A. Axelsson, and S.E. Holm. 1976. Studies on *Branhamella catarrhalis* (*Neisseria catarrhalis*) with special reference to maxillary sinusitis. Scan. J. Infect. Dis. 8:151-155.

2. Catlin, B.W., 1990. *Branhamella catarrhalis*: an organism gaining respect as a pathogen. Clin. Microbiol. Rev. 3: 293-320.

3. Hager, H., A. Verghese, S. Alvarez, and S.L. Berk. 1987. *Branhamella catarrhalis* respiratory infections. Rev. Infect. Dis. 9:1140-1149.

4. McLeod, D.T., F. Ahmad, M.J. Croughan, and M.A. Calder. 1986. Bronchopulmonary infection due to *M. catarrhalis*. Clinical features and therapeutic response. Drugs 31(Suppl.3):109-112.

5. Nicotra, B., M. Rivera, J.I. Luman, and R.J. Wallace. 1986. *Branhamella catarrhalis* as a lower respiratory tract pathogen in patients with chronic lung disease. Arch.Intern.Med. 146:890-893.

6. Ninane, G., J. Joly, and M. Kraytman. 1978. Bronchopulmonary infection due to *Branhamella catarrhalis* 11 cases assessed by transtracheal puncture. Br.Med.Jr. 1:276-278.

7. Srinivasan, G., M.J. Raff, W.C. Templeton, S.J. Givens, R.C. Graves, and J.C. Mel. 1981. *Branhamella catarrhalis* pneumonia. Report of two cases and review of the literature. Am.Rev. Respir. Dis. 123:553-555.

8. West, M., S.L. Berk, and J.K. Smith. 1982. *Branhamella catarrhalis* pneumonia. South.Med. J. 75:1021-1023.

9. Christensen, J.J., and B. Bruun. 1985. Bacteremia caused by a beta-lactamase producing strain of *Branhamella catarrhalis*. Acta.Pathol. Microbiol. Immunol. Scand. Sect.B 93:273-275.

10. Craig, D.B., and P.A. Wehrle. 1983. *Branhamella catarrhalis* septic arthritis. J. Rheumatol. 10:985-986.

11. Guthrie, R., K. Bakenhaster, R.Nelson, and R. Woskobnick. 1988. *Branhamella catarrhalis* sepsis: a

case report and review of the literature. J.Infect.Dis. 158:907-908.

12. Hiroshi, S., E.J. Anaissie, N.Khardori, and G.P. Bodey. 1988. *Branhamella catarrhalis* septicemia in patients with leukemia. Cancer 61:2315-2317.

13. O'Neill, J.H., and P.W. Mathieson. 1987. Meningitis due to *Branhamella catarrhalis*. Aust. N.Z. J. Med. 17:241-242.

14. Murphy, T.F. 1989. The surface of *Branhamella catarrhalis*: a systematic approach to the surface antigens of an emerging pathogen. Pediatr. Infect. Dis. J. 8:S75-S77.

15. Van Hare, G.F., P.A. Shurin, C.D. Marchant, N.A. Cartelli, C.E.Johnson, D. Fulton, S. Carlin, and C.H. Kim. Acute otitis media caused by *Branhamella catarrhalis*: biology and therapy. Rev. Infect. Dis. 9:16-27.

16. Jorgensen, J.H., Doern, G.V., Maher, L.A., Howell, A.W., and Redding, J.S., 1990 Antimicrobial resistance among respiratory isolates of *Haemophilus influenza, Moraxella catarrhalis*, and *Streptococcus* pneumoniae in the United States. Antibicrob. Agents Chemother. 34: 2075-2080.

17. Schryvers, A.B. and Morris, L.J. 1988 Identification and Characterization of the transferrin receptor from *Neisseria meningitidis*. Mol. Microbiol. 2:281-288.

18. Lee, B.C., Schryvers, A.B. Specificity of the lactoferrin and transferrin receptors in *Neisseria gonorrhoeae*. Mol. Microbiol. 1988; 2-827-9.

19. Schryvers, A.B. Characterization of the human transferrin and lactoferrin receptors in *Haemophilus influenzae*. Mol. Microbiol. 1988; 2: 467-72.

20. Schryvers, A.B. and Lee, B.C. (1988) Comparative analysis of the transferrin and lactoferrin binding proteins in the family *Neisseriaceae*. Can. J. Microbiol. 35, 409-415.

21. Yu, R. and Schryvers, A.B., 1993. The interaction between human transferrin and transferrin binding protein 2 from *Moraxella (Branhamella) catarrhalis*

differs from that of other human pathogens. Microbiol. Pathogenesis, 15:433-445.

22. O'Hagan, 1992. Clin. Pharmokinet. 22:1.

23. Ulmer et al., 1993. Curr. Opinion Invest. Drugs 2: 983-989.

24. Lockhoff, O., 1991. glycolipds as immunomoclutators: Synthesis and properits. Chem. Int. Ed. Engl. 30: 1611-1620.

25. Nixon-George, 1990. J. Immunol. 14: 4798-4802.

26. Wallace, R.J. Jr., Nash, D.R., and Steingrube, V.A. 1990. Antibiotic susceptibilites and drug resistance in *Moraxella (Branhaemella) catarrhalis*. Am. J. Med. 88 (5A): 465-50S.

27. F.M. Ausubel et al., Short protocols in Molecular Biology, Greene Publishing Associates and John Wiley and Sons.

28. Legrain, M., V. Mazarin, S.W. Irwin, B. Bouchon, M-J. Quentin-Millet, E. Jacobs, and A.B. Schryvers. 1993, Cloning and characterization of Neisseria meningitidis genes encoding the transferrin-binding proteins Tbpl and Tbp2. Gene 130: 73-80.

29. Ogunnariwo, J.W., Woo, T.K.W., Lo, R.Y.C., Gonzalez, G.C., and Schryvers, A.B. Characterization of the Pasteurella haemolytica transferrin receptor genes and the recombinant receptor proteins. Microb. Pathog. 23:273-284 (1997).

30. Yang, Y.P., Myers, L.E., McGuinness, U., Chong, P., Kwok, Y., Klein, M.H. and Harkness R.E. The major outer membrane protein, C.D, extracted from Moraxella (Branhamella) catarrhalis is a potential vaccine antigen that induces bactericidal antibodies. FEMS Immun. Med. Microbiol. 17:187-199 (1997).

## CLAIMS

What we claim is:

1. A purified and isolated nucleic acid molecule encoding a Tbp2 protein of a strain of *Moraxella* which strain is selected from the group consisting of *Moraxella catarrhalis* M35, 3 and LES1.

2. The purified and isolated nucleic acid molecule of claim 1, having a DNA sequence selected from the group consisting of:

    (a) a DNA sequence as set out in Figure 2, 4 or 6 (SEQ ID NOS: 1, 3 or 5) or the complementary DNA sequence thereto; or

    (b) a DNA sequence encoding an amino acid sequence as set out in Figure 2, 4 or 6 (SEQ ID NOS: 2, 4 or 6) or the complementary DNA sequence thereto.

3. A vector adapted for transformation of a host comprising the nucleic acid molecule of claim 1.

4. The vector of claim 3 further comprising expression means operatively coupled to the nucleic acid molecule for expression by the host of said Tbp2 protein of a *Moraxella catarrhalis* strain M35, 3 or LES1.

5. A transformed host containing an expression vector as claimed in claim 4.

6. A method of forming a substantially pure recombinant Tbp2 protein of a *Moraxella catarrhalis* strain M35, 3 or LES1 which comprises:

growing the transformed host of claim 5 to express Tbp2 protein as inclusion bodies,

purifying the inclusion bodies free from cellular material and soluble proteins,

solubilizing Tbp2 protein from the purified inclusion bodies, and

purifying the Tbp2 protein free from other solubilized materials.

7. A recombinant Tbp2 protein of *Moraxella catarrhalis* strain M35, 3 or LES1 producible by the transformed host of claim 5, having a deduced amino acid sequence selected from the group consisting of those shown in Figure 2, 4 or 6 (SEQ ID NO: 2, 4 or 6).

8. An immunogenic composition, comprising at least one active component selected from the group consisting of:

(A) a purified and isolated nucleic acid molecule as claimed in claim 1; or

(B) a recombinant Tbp2 protein as claimed in claim 7;

and a pharmaceutically acceptable carrier therefor, said at least one active component producing an immune response when administered to a host.

9. A method for generating an immune response in a host, comprising administering to the host an immunoeffective amount of the immunogenic composition of claim 8.

10. A method of determining the presence, in a sample, of nucleic acid encoding a transferrin receptor protein of a strain of *Moraxella*, comprising the steps of:

(a) contacting the sample with the nucleic acid molecule of claim 1 to produce duplexes comprising the nucleic acid molecule and any said nucleic acid molecule encoding the transferrin receptor protein of a strain of *Moraxella* present in the sample and specifically hybridizable therewith; and

(b) determining production of the duplexes.

11. A diagnostic kit for determining the presence, in a sample, of nucleic acid encoding a transferrin receptor protein of a strain of *Moraxella*, comprising:

(a) the nucleic acid molecule of claim 1;

(b) means for contacting the nucleic acid molecule with the sample to produce duplexes comprising the nucleic acid molecule and any said nucleic acid present in the sample and hybridizable with the nucleic acid molecule; and

(c) means for determining production of the duplexes.

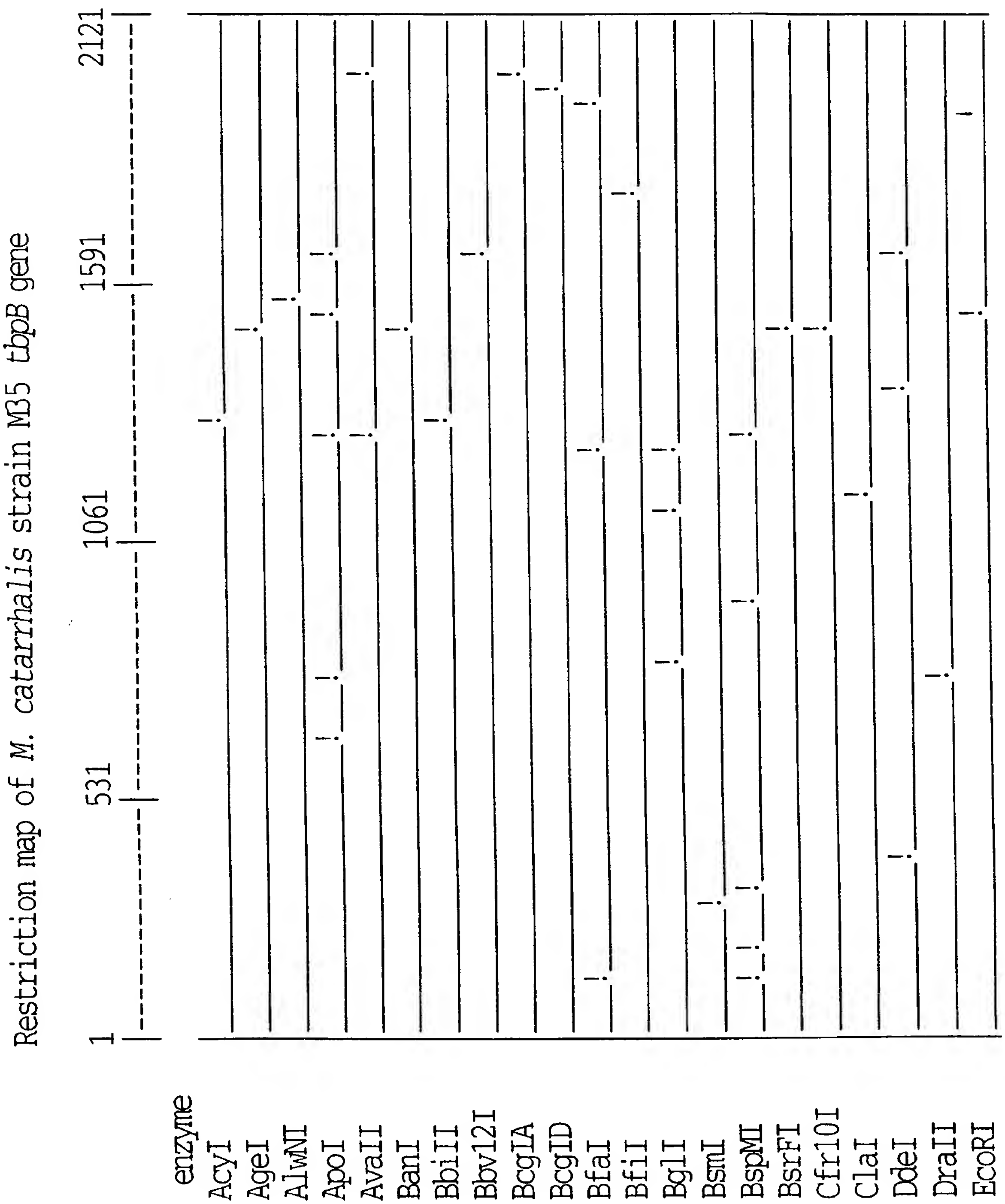
12. A nucleic acid molecule of claim 1 when used as a medicine.

13. A recombinant transferrin receptor protein of claim 7 when used as a medicine.

14. The use of a nucleic acid molecule of claim 1 in the manufacture of a medicament for protection against infection by a strain of *Moraxella*.

15. The use of a recombinant transferrin receptor protein of claim 7 in the manufacture of a medicament for protection against infection by a strain of *Moraxella*.

FIG.1A


Restriction map of *M. catarrhalis* strain M35 *tbpB* gene
1
531
1061
1591
2121
enzyme
AcyI
AgeI
AlwNI
ApoI
AvaII
BanI
BbiII
Bbv12I
BcgIA
BcgID
BfaI
BfiI
BglI
BsmI
BspMI
BsrFI
Cfr10I
ClaI
DdeI
DraII
EcoRI

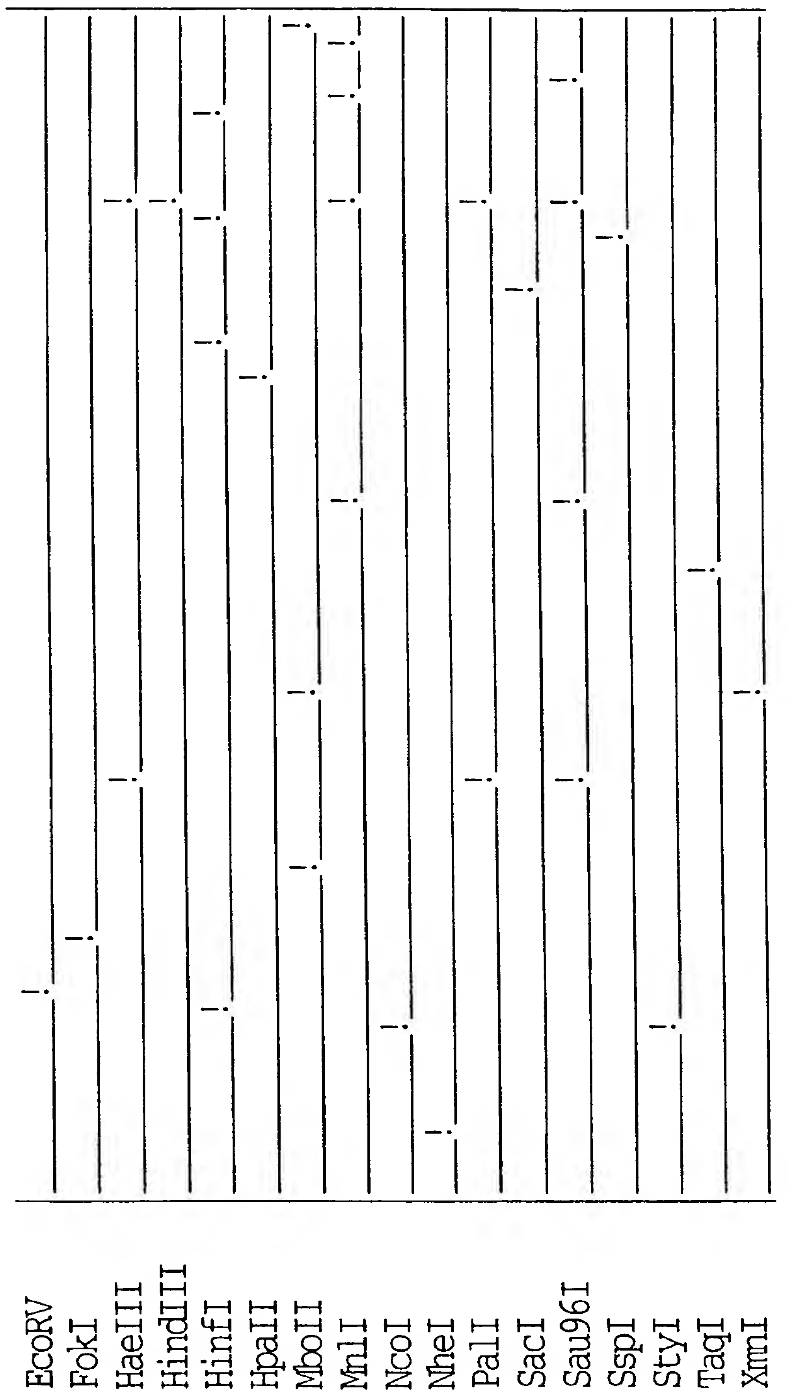

EcoRV
FokI
HaeIII
HindIII
HinfI
HpaII
MboII
MnlI
NcoI
NheI
PalI
SacI
Sau96I
SspI
StyI
TaqI
XmnI

FIG.1B

# FIG.2A

*M. catarrhalis* strain M35 *tbpB* sequence

```
MET   LYS   HIS   ILE   PRO   LEU   THR   THR   LEU   CYS ...
A T G A A A C A C A T T C C T T T A A C C A C A C T G T G T...
                  10                  20                  30...
                                ...  VAL   ALA   ILE   SER   ALA   VAL   LEU   LEU   THR   ALA
                                ... G T G G C A A T C T C T G C C G T C T T A T T A A C C G C T
                                ...                   40                  50                  60

CYS   GLY   GLY   SER   GLY   GLY   SER   ASN   PRO   PRO ...
T G T G G T G G C A G T G G T G G T T C A A A T C C A C C T...
                  70                  80                  90...
                                ...  ALA   PRO   THR   PRO   ILE   PRO   ASN   ALA   SER   GLY
                                ... G C T C C T A C G C C C A T T C C A A A T G C T A G C G G T
                                ...                  100                 110                 120

SER   GLY   ASN   THR   GLY   ASN   THR   GLY   ASN   ALA ...
T C A G G T A A T A C T G G C A A C A C T G G T A A T G C T...
                 130                 140                 150...
                                ...  GLY   GLY   THR   ASP   ASN   THR   ALA   ASN   ALA   GLY
                                ... G G C G G T A C T G A T A A T A C A G C C A A T G C A G G T
                                ...                  160                 170                 180

ASN   THR   GLY   GLY   THR   ASN   SER   GLY   THR   GLY ...
A A T A C A G G C G G T A C A A A C T C T G G T A C A G G C...
                 190                 200                 210...
                                ...  SER   ALA   ASN   THR   PRO   GLU   PRO   LYS   TYR   LYS
                                ... A G T G C C A A C A C A C C A G A A C C A A A A T A T A A A
                                ...                  220                 230                 240
```

FIG.2B

```
ASP  VAL  PRO  THR  ASP  GLU  ASN  LYS  LYS  ASP ...
G A T G T G C C A A C C G A T G A A A A T A A A A A A G A T...
               250                 260                 270...
                                            ... GLU  VAL  SER  GLY  ILE  GLN  GLU  PRO  ALA  MET
                                            ... G A A G T G T C A G G C A T T C A A G A A C C T G C C A T G
                                            ...                280                 290                 300

GLY  TYR  GLY  MET  ALA  LEU  SER  LYS  MET  ASN ...
G G T T A T G G C A T G G C T T T G A G T A A A A T G A A T...
               310                 320                 330...
                                            ... LEU  HIS  LYS  GLN  GLN  ASP  THR  PRO  LEU  ASP
                                            ... C T A C A C A A A C A A C A A G A C A C G C C A T T A G A T
                                            ...                340                 350                 360

GLU  LYS  ASP  ILE  ILE  THR  LEU  ASP  GLY  LYS ...
G A A A A A G A T A T C A T T A C C T T A G A C G G T A A A...
               370                 380                 390...
                                            ... LYS  GLN  VAL  ALA  LYS  GLY  GLU  LYS  SER  PRO
                                            ... A A A C A A G T T G C A A A A G G T G A A A A A T C G C C A
                                            ...                400                 410                 420

LEU  PRO  PHE  SER  LEU  ASP  VAL  GLU  ASN  LYS ...
T T G C C A T T T T C G T T G G A T G T A G A A A A T A A A...
               430                 440                 450...
                                            ... LEU  LEU  ASP  GLY  TYR  ILE  ALA  LYS  MET  ASN
                                            ... T T G C T T G A T G G C T A T A T A G C A A A A A T G A A T
                                            ...                460                 470                 480
```

FIG.2C

```
GLU ALA ASP LYS ASN ALA ILE GLY ASP ARG ...
G A A G C G G A T A A A A A T G C C A T T G G T G A C A G A...
490 500 510...
... ILE LYS LYS ASP ASN LYS ASP LYS SER LEU
... A T T A A G A A A G A T A A T A A A G A C A A G T C A T T A
... 520 530 540

SER LYS ALA GLU LEU ALA LYS GLN ILE LYS ...
T C T A A A G C A G A G C T T G C C A A A C A A A T C A A A...
550 560 570...
... GLU ASP VAL ARG LYS SER HIS GLU PHE GLN
... G A A G A T G T G C G T A A A A G C C A T G A G T T T C A G
... 580 590 600

GLN VAL LEU SER SER LEU LYS ASN LYS ILE ...
C A A G T A T T A T C A T C A C T G A A A A A C A A A A T T...
610 620 630...
... PHE HIS SER ASN ASP GLY THR THR LYS ALA
... T T T C A T T C A A A T G A T G G A A C A A C C A A A G C A
... 640 650 660

THR THR ARG ASP LEU GLN TYR VAL ASP TYR
A C C A C A C G A G A T T T A C A A T A T G T T G A T T A T
670 680 690
... GLY TYR TYR LEU VAL ASN ASP GLY ASN TYR
... G G T T A C T A C T T G G T G A A T G A T G G C A A T T A T
... 700 710 720
```

FIG.2D

```
LEU   THR   VAL   LYS   THR   ASP   GLU   LEU   TRP   ASN ...
C T A A C C G T C A A A A C A G A C G A A C T T T G G A A T
              730                     740                 750
                                ... LEU   GLY   PRO   VAL   GLY   GLY   VAL   PHE   TYR   ASN
                                ... T T A G G C C C T G T G G G C G G T G T G T T T T A T A A T
                                ...               760                     770                780

GLY   THR   THR   THR   ALA   LYS   GLU   LEU   PRO   THR ...
G G C A C A A C G A C C G C C A A A G A G C T A C C C A C A...
              790                     800                 810...
                                ... GLN   ASP   ALA   VAL   LYS   TYR   LYS   GLY   HIS   TRP
                                ... C A A G A T G C G G T C A A A T A T A A A G G A C A T T G G
                                ...               820                     830                840

ASP   PHE   MET   THR   ASP   VAL   ALA   LYS   GLN   ARG ...
G A C T T T A T G A C C G A T G T T G C C A A A C A A A G A...
              850                     860                 870...
                                ... ASN   ARG   PHE   SER   GLU   VAL   LYS   GLU   ASN   LEU
                                ... A A C C G A T T T A G C G A A G T G A A A G A A A A C C T T
                                ...               880                     890                900

GLN   ALA   GLY   ARG   TYR   TYR   GLY   ALA   SER   SER ...
C A A G C A G G T C G G T A T T A T G G A G C A T C T T C A...
              910                     920                 930...
                                ... LYS   ASP   GLU   TYR   ASN   ARG   LEU   LEU   THR   ASP
                                ... A A A G A T G A A T A C A A C C G C T T A T T A A C T G A T
                                ...               940                     950                960
```

FIG.2E

```
GLU   LYS   ASN   LYS   PRO   GLU   ARG   TYR   ASN   GLY ...
G A G A A A A A C A A A C C A G A G C G T T A T A A C G G T...
              970                 980                 990...
                                ... GLU   TYR   GLY   HIS   SER   SER   GLU   PHE   THR   VAL
                                ... G A A T A T G G T C A T A G C A G T G A G T T T A C T G T T
                                ...               1000                1010                1020

ASN   PHE   LYS   ASP   LYS   LYS   LEU   THR   GLY   GLU ...
A A T T T T A A G G A C A A A A A A T T A A C A G G T G A G...
              1030                1040                1050
                                ... LEU   PHE   SER   ASN   LEU   GLN   ASP   SER   ARG   LYS
                                ... C T G T T T A G T A A C C T A C A A G A C A G C C G T A A G
                                ...               1060                1070                1080

GLY   ASN   VAL   THR   LYS   THR   LYS   ARG   TYR   ASP ...
G G C A A T G T T A C G A A A A C C A A A C G C T A T G A C...
              1090                1100                1110...
                                ... ILE   ASP   ALA   ASN   ILE   TYR   GLY   ASN   ARG   PHE
                                ... A T C G A T G C C A A T A T C T A C G G C A A C C G C T T C
                                              1120                1130                1140

ARG   GLY   SER   ALA   THR   ALA   SER   ASP   LYS   ALA ...
C G T G G C A G T G C C A C C G C A A G C G A T A A A G C A...
              1150                1160                1170...
                                ... GLU   ALA   SER   LYS   THR   LYS   HIS   PRO   PHE   THR
                                ... G A A G C A A G C A A A A C C A A A C A C C C C T T T A C C
                                ...               1180                1190                1200
```

FIG.2F

```
SER  ASP  ALA  LYS  ASN  SER  LEU  GLU  GLY  GLY ...
A G C G A T G C C A A A A A T A G C C T A G A A G G C G G T...
          1210                1220                1230...
                    ... PHE  TYR  GLY  PRO  ASN  ALA  GLU  GLU  LEU  ALA
                    ... T T T T A T G G A C C A A A C G C C G A G G A G C T G G C A
                    ...           1240                1250                1260

GLY  LYS  PHE  LEU  THR  ASN  ASP  ASN  LYS  LEU ...
G G T A A A T T C C T A A C C A A T G A C A A C A A A C T C...
          1270                1280                1290...
                    ... PHE  GLY  VAL  PHE  GLY  ALA  LYS  ARG  GLU  SER
                    ... T T T G G C G T C T T T G G T G C T A A A C G A G A G T
                    ...           1300                1310                1320

LYS  ALA  GLY  GLU  LYS  THR  GLU  ALA  ILE  LEU ...
A A A G C T G G G G A A A A A A C C G A A G C C A T C T T A...
          1330                1340                1350
                    ... ASP  ALA  TYR  ALA  LEU  GLY  THR  PHE  ASN  LYS
                    ... G A T G C C T A T G C A C T T G G G A C A T T T A A C A A A
                    ...           1360                1370                1380

ASN  ASN  ALA  THR  THR  PHE  THR  PRO  PHE  THR ...
A A T A A C G C A A C C A C A T T C A C C C C A T T T A C C...
          1390                1400                1410...
                    ... LYS  LYS  GLN  LEU  ASP  ASN  PHE  GLY  ASN  ALA
                    ... A A A A A A C A A C T G G A T A A C T T T G G C A A T G C C
                    ...           1420                1430                1440
```

FIG.2G

```
LYS   LYS   LEU   VAL   LEU   GLY   SER   THR   VAL   ILE ...
A A A A A G T T G G T C T T G G G T T C T A C C G T C A T T...
          1450                1460                1470
                        ... ASP   LEU   VAL   PRO   THR   GLY   VAL   THR   LYS   ASP
                        ... G A T T T G G T G C C T A C C G G T G T C A C C A A A G A T
                        ...                1480                1490                1500

VAL   ASN   GLU   PHE   THR   LYS   ASN   LYS   PRO   ASP ...
G T C A A T G A A T T C A C C A A A A A C A A G C C A G A T...
          1510                1520                1530
                        ... SER   ALA   THR   ASN   LYS   ALA   GLY   GLU   THR   LEU
                        ... T C T G C C A C A A A C A A A G C G G G C G A G A C T T T G
                        ...                1540                1550                1560

MET   VAL   ASN   ASP   LYS   VAL   SER   VAL   LYS   THR ...
A T G G T G A A T G A T A A A G T T A G C G T C A A A A C C...
          1570                1580                1590...
                        ... TYR   GLY   TYR   GLY   ARG   ASN   PHE   GLU   TYR   LEU
                        ... T A T G G C T A T G G C A G A A A C T T T G A A T A C C T A
                        ...                1600                1610                1620

LYS   PHE   GLY   GLU   LEU   SER   VAL   GLY   THR   SER ...
A A A T T T G G T G A G C T C A G T G T C G G C A C A A G C...
          1630                1640                1650...
                        ... ASN   SER   VAL   PHE   LEU   GLN   GLY   GLU   ARG   THR
                        ... A A C A G C G T C T T T T T A C A A G G C G A A C G C A C C
                        ...                1660                1670                1680
```

FIG.2H

```
ALA THR THR GLY GLU LYS ALA VAL PRO THR ...
G C T A C C A C A G G C G A G A A A G C C G T A C C A A C C...
        1690            1700             1710...
... LYS GLY THR ALA LYS TYR LEU GLY ASN TRP
... A A A G G C A C A G C C A A A T A T T T G G G G A A C T G G
...             1720            1730            1740

VAL GLY TYR ILE THR GLY LYS ASP SER SER ...
G T A G G A T A C A T C A C A G G A A A G G A C T C A T C A...
        1750            1760             1770...
... LYS SER PHE ASN GLU ALA GLN ASP VAL ALA
... A A A A G C T T T A A T G A G G C C C A A G A T G T T G C T
...             1780            1790            1800

ASP PHE ASP ILE ASP PHE GLU LYS LYS SER ...
G A T T T T G A C A T T G A C T T T G A G A A A A A A T C A...
        1810            1820             1830...
... VAL LYS GLY LYS LEU THR THR LYS ASP ARG
... G T T A A A G G C A A A C T G A C C A C C A A A G A C C G C
...             1840            1850            1860

GLN ASP PRO VAL PHE ASN ILE THR GLY ASP ...
C A A G A C C C T G T A T T T A A C A T C A C A G G T G A C...
        1870            1880             1890...
... ILE ALA GLY ASN GLY TRP THR GLY LYS ALA
... A T C G C A G G C A A T G G C T G G A C A G G C A A A G C C
...             1900            1910            1920
```

FIG.2I

```
SER THR THR LYS ALA ASP ALA GLY GLY TYR ...
A G C A C C A C C A A A G C G G A C G C A G G G G G C T A C...
        1930                1940                1950...
                            ... LYS ILE ASP SER SER SER THR GLY LYS SER
                            ... A A G A T A G A T T C T A G C A G T A C A G G C A A A T C C
                            ...         1960                1970                1980

ILE VAL ILE LYS ASP ALA GLU VAL THR GLY ...
A T C G T C A T C A A A G A T G C C G A G G T T A C A G G G...
        1990                2000                2010...
                            ... GLY PHE TYR GLY PRO ASN ALA ASN GLU MET
                            ... G G C T T T T A T G G T C C A A A T G C A A A C G A G A T G
                            ...         2020                2030                2040

GLY GLY SER PHE THR HIS ASN THR ASP ASP ...
G G C G G G T C A T T T A C A C A C A A C A C C G A T G A C...
        2050                2060                2070...
                            ... SER LYS ALA SER VAL VAL PHE GLY THR LYS
                            ... A G T A A A G C C T C T G T G G T C T T T G G C A C A A A A
                            ...         2080                2090                2100
ARG GLN GLU GLU VAL LYS ***
A G A C A A G A A G A A G T T A A G T A G
        2110                2120
```

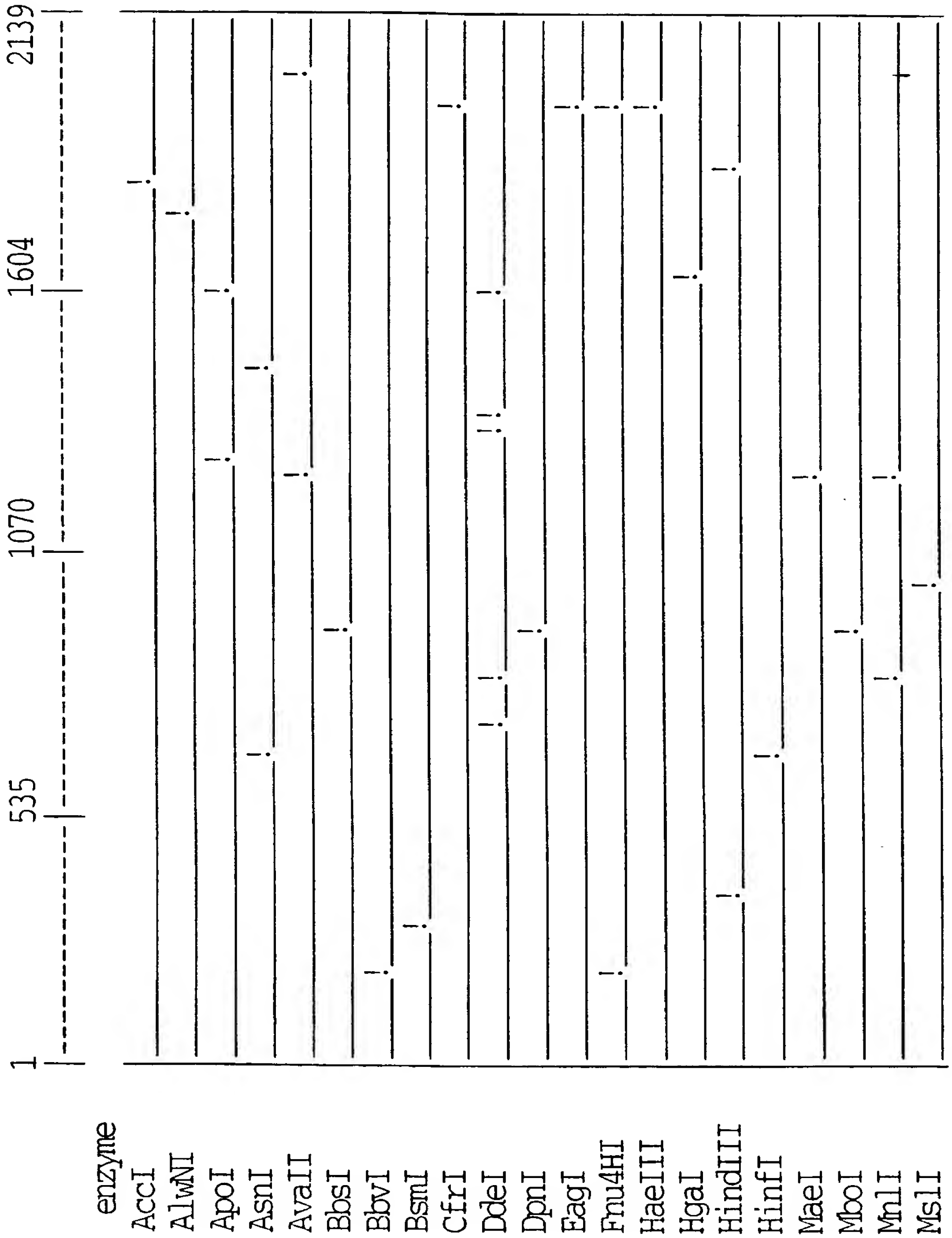



FIG.3A

Restriction map of *M. catarrhalis* strain 3 *tbpB* gene

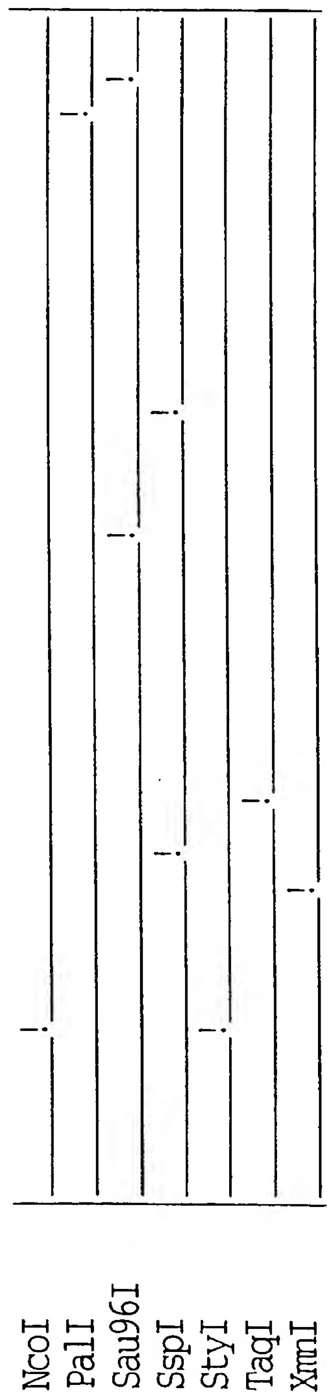



FIG.3B

# FIG.4A

*M. catarrhalis* strain 3 *tbpB* sequence

```
MET LYS HIS ILE PRO LEU THR THR LEU CYS ...
A T G A A A C A C A T T C C T T T A A C C A C A C T G T G T...
10 20 30...
... VAL ALA ILE SER ALA VAL LEU LEU THR ALA
... G T G G C A A T C T C T G C C G T C T T A T T A A C C G C T
... 40 50 60

CYS GLY GLY SER GLY GLY SER ASN PRO PRO ...
T G T G G T G G C A G T G G T G G T T C A A A T C C A C C T...
70 80 90...
... ALA PRO THR PRO ILE PRO ASN ALA GLY GLY
... G C T C C T A C G C C C A T T C C A A A T G C A G G C G G T
... 100 110 120

ALA GLY ASN ALA GLY SER GLY THR GLY GLY ...
G C A G G T A A T G C T G G T A G C G G T A C T G G C G G T...
130 140 150...
... ALA GLY SER THR ASP ASN ALA ALA ASN ALA
... G C A G G T A G C A C T G A T A A T G C A G C C A A T G C A
... 160 170 180

GLY SER THR GLY GLY ALA SER SER GLY THR ...
G G C A G T A C A G G C G G T G C A A G C T C T G G T A C A...
190 200 210...
... GLY SER ALA SER THR GLN LYS PRO LYS TYR
... G G C A G T G C C A G C A C A C A A A A A C C A A A A T A T
... 220 230 240
```

FIG.4B

```
GLN   ASP   VAL   PRO   THR   ASP   LYS   ASN   LYS   LYS ...
C A A G A T G T G C C A A C C G A T A A A A A T A A A A A A...
              250                 260                 270...
                                        ...  ASP   GLU   VAL   SER   GLY   ILE   GLN   GLU   PRO   ALA
                                        ... G A T G A A G T G T C A G G C A T T C A A G A A C C T G C C
                                        ...                 280                 290                 300

MET   GLY   TYR   GLY   VAL   GLU   LEU   LYS   LEU   ARG ...
A T G G G T T A T G G C G T G G A A T T A A A G C T T C G T...
              310                 320                 330...
                                        ...  ASN   TRP   ILE   PRO   GLN   GLU   GLN   GLU   GLU   HIS
                                        ... A A C T G G A T A C C A C A A G A A C A G G A A G A A C A T
                                        ...                 340                 350                 360

ALA   LYS   ILE   ASN   THR   ASN   ASP   VAL   VAL   LYS ...
G C C A A A A T C A A T A C A A A T G A T G T T G T A A A A...
              370                 380                 390...
                                        ...  LEU   GLU   GLY   ASP         LEU   LYS   HIS   ASN   PRO   PHE
                                        ... C T T G A A G G T G A C T T G A A G C A T A A T C C A T T T
                                        ...                 400                 410                 420

ASP   ASN   SER   ILE   TRP   GLN   ASN   ILE   LYS   ASN ...
G A C A A C T C T A T T T G G C A A A A C A T C A A A A A T...
              430                 440                 450...
                                        ...  SER   LYS   GLU   VAL   GLN   THR   VAL   TYR   ASN   GLN
                                        ... A G C A A A G A A G T A C A A A C T G T T T A C A A C C A A
                                        ...                 460                 470                 480
```

FIG.4C

```
GLU  LYS  GLN  ASN  ILE  GLU  ASN  GLN  ILE  LYS ...
G A G A A G C A A A A C A T T G A A A A T C A A A T C A A A...
            490                 500                 510...
                                 ... LYS  GLU  ASN  LYS  GLU  LEU  ASP  LYS  THR  ALA
                                 ... A A A G A A A A T A A A G A A C T T G A T A A A A C G G C A
                                 ...                520                 530                 540

LEU  LYS  ALA  LEU  ILE  GLU  LYS  VAL  LEU  ASP ...
C T A A A A G C T C T T A T T G A A A A A G T T C T T G A T...
            550                 560                 570...
                                 ... ASP  TYR  LEU  THR  SER  LEU  ALA  LYS  PRO  ILE
                                 ... G A C T A T C T A A C A A G T C T T G C T A A A C C C A T T
                                 ...                580                 590                 600

TYR  GLU  LYS  ASN  ILE  ASN  ASP  SER  HIS  ASP ...
T A T G A A A A A A A T A T T A A T G A T T C A C A T G A T...
            610                 620                 630...
                                 ... LYS  GLN  ASN  LYS  ALA  ARG  THR  ARG  ASP  LEU
                                 ... A A G C A G A A T A A A G C A C G C A C T C G T G A T T T G
                                 ...                640                 650                 660

LYS  TYR  VAL  ARG  SER  GLY  TYR  ILE  TYR  ARG ...
A A G T A T G T G C G T T C T G G T T A T A T T T A T C G C...
            670                 680                 690...
                                 ... SER  GLY  TYR  SER  ASN  ILE  ASP  ILE  GLN  LYS
                                 ... T C A G G T T A T T C T A A T A T C G A C A T T C A A A A G
                                 ...                700                 710                 720
```

FIG.4D

```
LYS  ILE  ALA  LYS  THR  GLY  PHE  ASP  GLY  ALA ...
A A A A T A G C T A A A A C T G G T T T T G A T G G T G C T...
          730                 740                 750...
                               ... LEU  PHE  TYR  LYS  GLY  THR  GLN  THR  ALA  LYS
                               ... T T A T T T T A T A A A G G T A C A C A A A C T G C T A A A
                               ...                760                 770                 780

GLN  LEU  PRO  VAL  SER  GLU  VAL  LYS  TYR  LYS ...
C A A T T G C C T G T A T C T G A G G T T A A G T A T A A A...
          790                 800                 810...
                               ... GLY  THR  TRP  ASP  PHE  MET  THR  ASP  ALA  LYS
                               ... G G C A C T T G G G A T T T T A T G A C C G A T G C C A A A
                               ...                820                 830                 840

LYS  GLY  GLN  SER  PHE  SER  SER  PHE  GLU  ARG ...
A A A G G A C A A T C A T T T A G C A G T T T T G A A A G A...
          850                 860                 870...
                               ... ARG  ALA  GLY  ASP  ARG  TYR  SER  ALA  MET  SER
                               ... C G A G C T G G T G A T C G C T A T A G T G C A A T G T C T
                               ...                880                 890                 900

SER  HIS  GLU  TYR  PRO  SER  LEU  LEU  THR  ASP ...
T C C C A T G A G T A C C C A T C T T T A T T A A C T G A T...
          910                 920                 930...
                               ... ASP  LYS  ASN  LYS  PRO  ASP  ASN  TYR  ASN  ASP
                               ... G A T A A A A A C A A A C C A G A T A A T T A T A A C G A T
                               ...                940                 950                 960
```

FIG.4E

```
GLU TYR GLY HIS SER SER GLU PHE THR VAL ...
G A A T A T G G T C A T A G C A G T G A G T T T A C G G T A...
        970                    980                    990...
... ASP PHE SER LYS LYS SER LEU THR GLY GLY
... G A T T T T A G T A A A A A G A G C C T A A C A G G T G G G
...                1000                  1010                  1020

LEU PHE SER ASN LEU GLN ASP HIS HIS LYS ...
C T G T T T A G T A A C C T A C A A G A C C A C C A T A A G...
        1030                   1040                   1050...
... GLY LYS VAL THR LYS THR LYS ARG TYR ASP
... G G C A A G G T T A C G A A A A C C A A A C G C T A T G A C
...                1060                  1070                  1080

ILE ASN ALA ARG ILE HIS GLY ASN ARG PHE ...
A T C A A T G C C C G T A T C C A C G G T A A C C G C T T C...
        1090                   1100                   1110...
... ARG GLY SER ALA THR ALA ILE ASN LYS ASP
... C G T G G C A G T G C C A C C G C A A T C A A T A A A G A T
...                1120                  1130                  1140

ASN GLU SER LYS ALA LYS HIS PRO PHE THR ...
A A T G A A A G C A A A G C C A A A C A C C C C T T T A C C...
        1150                   1160                   1170...
... SER ASP ALA ASP ASN ARG LEU GLU GLY GLY
... A G C G A T G C C G A C A A T A G G C T A G A A G G C G G T
...                1180                  1190                  1200
```

# FIG.4F

```
PHE   TYR   GLY   PRO   ASN   ALA   GLU   GLU   LEU   ALA ...
T T T T A T G G A C C A A A C G C C G A G G A G C T G G C A...
               1210                1220                1230...
                                ... GLY   LYS   PHE   LEU   THR   ASP   ASP   ASN   LYS   LEU
                                ... G G T A A A T T C C T A A C C G A T G A C A A C A A A C T C
                                ...                1240                1250                1260

PHE   GLY   VAL   PHE   GLY   ALA   LYS   GLN   GLU   SER ...
T T T G G T G T C T T T G G T G C T A A A C A A G A G A G T...
               1270                1280                1290...
                                ... GLU   ALA   LYS   GLU   THR   GLU   ALA   ILE   LEU   ASP
                                ... G A A G C T A A G G A A A C C G A A G C C A T C T T A G A T
                                ...                1300                1310                1320

ALA   TYR   ALA   LEU   GLY   THR   PHE   ASN   LYS   SER ...
G C T T A T G C A C T T G G G A C A T T T A A T A A A T C T...
               1330                1340                1350...
                                ... GLY   THR   THR   ASN   PRO   ALA   PHE   THR   ALA   ASN
                                ... G G T A C G A C C A A T C C T G C C T T T A C C G C C A A T
                                ...                1360                1370                1380

SER   LYS   LYS   GLU   LEU   ASP   ASN   PHE   GLY   ASN ...
A G T A A A A A A G A A C T G G A T A A C T T T G G C A A T...
               1390                1400                1410...
                                ... ILE   ASN   LYS   LEU   VAL   LEU   GLY   SER   THR   VAL
                                ... A T T A A T A A A T T G G T C T T G G G T T C T A C T G T G
                                ...                1420                1430                1440
```

FIG.4G

```
ILE   ASP   LEU   THR   GLN   GLY   ASN   ASP   PHE   VAL ...
A T A G A C C T T A C T C A A G G T A A T G A T T T T G T A...
             1450                1460                1470...
                                     ... LYS   THR   ILE   ASP   LYS   GLU   LYS   PRO   ALA   THR
                                     ... A A A A C C A T T G A T A A A G A A A A G C C A G C C A C C
                                     ...            1480                1490                1500

THR   THR   ASN   GLN   ALA   GLY   GLU   PRO   LEU   THR ...
A C T A C C A A T C A A G C A G G C G A G C C T T T G A C G...
             1510                1520                1530...
                                     ... VAL   ASN   ASP   LYS   VAL   ARG   VAL   GLN   VAL   CYS
                                     ... G T G A A T G A T A A G G T T C G G G T A C A A G T T T G T
                                     ...            1540                1550                1560

CYS   SER   ASN   LEU   GLU   HIS   LEU   LYS   PHE   GLY ...
T G T A G C A A T C T T G A G C A T C T A A A A T T T G G C...
             1570                1580                1590...
                                     ... SER   LEU   SER   ILE   GLY   ASP   SER   ASN   SER   VAL
                                     ... T C A C T G A G T A T C G G T G A T A G T A A T A G C G T C
                                         ...        1600                1610                1620

PHE   LEU   GLN   GLY   GLU   ARG   THR   ALA   THR   LYS ...
T T T T T A C A A G G T G A A C G C A C C G C T A C C A A A...
             1630                1640                1650...
                                     ... GLY   ASP   LYS   ASP   LYS   ALA   MET   PRO   VAL   ALA
                                     ... G G T G A T A A A G A T A A A G C C A T G C C A G T T G C A
                                     ...            1660                1670                1680
```

# FIG.4H

```
GLY   ASN   ALA   LYS   TYR   ARG   GLY   THR   TRP   ALA ...
G G A A A T G C T A A A T A C C G T G G T A C A T G G G C A...
                 1690                1700                1710...
                              ... GLY   TYR   VAL   ALA   GLY   SER   GLY   ASN   THR   SER
                              ... G G C T A T G T T G C A G G C T C T G G C A A T A C C A G C
                              ...                1720                1730                1740

LYS   ALA   TYR   GLU   ALA   GLN   GLN   PHE   ALA   ASP ...
A A A G C C T A T G A A G C C C A A C A A T T T G C T G A C...
                 1750                1760                1770...
                              ... ASN   ALA   ASN   ARG   ALA   GLU   PHE   ASP   VAL   ASP
                              ... A A T G C C A A C C G T G C C G A G T T T G A T G T A G A C
                              ...                1780                1790                1800

PHE   ALA   ASN   LYS   SER   LEU   THR   GLY   LYS   LEU ...
T T T G C T A A C A A A A G C C T A A C T G G T A A G C T T...
                 1810                1820                1830...
                              ... ILE   PRO   ASN   THR   SER   SER   ASP   GLY   LYS   SER
                              ... A T T C C A A A T A C G A G C A G T G A T G G T A A A T C T
                              ...                1840                1850                1860

ALA   PHE   ASP   ILE   THR   ALA   THR   ILE   ASP   GLY ...
G C T T T T G A T A T T A C T G C T A C A A T T G A T G G C...
                 1870                1880                1890...
                              ... ASN   GLY   PHE   SER   GLY   LYS   ALA   ASN   THR   PRO
                              ... A A T G G T T T T A G T G G T A A A G C C A A T A C A C C A
                              ...                1900                1910                1920
```

FIG.4I

```
ASP  ILE  GLU  THR  GLY  GLY  LEU  LYS  ILE  ASP ...
G A T A T T G A A A C A G G T G G G T T A A A G A T T G A C...
            1930            1940           1950...
                    ... SER  LYS  ASN  SER  GLU  SER  GLY  ARG  VAL  ILE
                    ... A G T A A G A A C A G T G A A A G C G G C C G A G T A A T T
                    ...            1960           1970           1980

VAL  LYS  ASP  ALA  ILE  VAL  ILE  GLY  GLY  PHE ...
G T G A A A G A T G C T A T A G T T A T A G G T G G C T T T...
            1990            2000           2010...
                    ... TYR  GLY  PRO  GLN  ALA  ASN  GLU  LEU  GLY  GLY
                    ... T A T G G T C C A C A A G C T A A T G A A C T G G G T G G C
                    ...            2020           2030           2040

SER  PHE  THR  TYR  LYS  SER  ASN  ASP  ALA  GLY ...
T C A T T T A C C T A C A A G A G C A A T G A T G C T G G A...
            2050            2060           2070...
                    ... ASN  GLN  ASP  LYS  ASP  SER  SER  ALA  SER  VAL
                    ... A A T C A A G A C A A A G A C A G T A G T G C A T C T G T G
                    ...            2080           2090           2100

VAL  PHE  GLY  ALA  ARG  LYS  GLN  GLN  GLU  VAL ...
G T C T T T G G T G C A A G A A A A C A A C A A G A A G T C...
            2110            2120           2130
                    ... LYS  PRO  ***
                    ... A A A C C A T G A
```

FIG.5A

Restriction map of *M. catarrhalis* strain LES1 *tbpB* gene

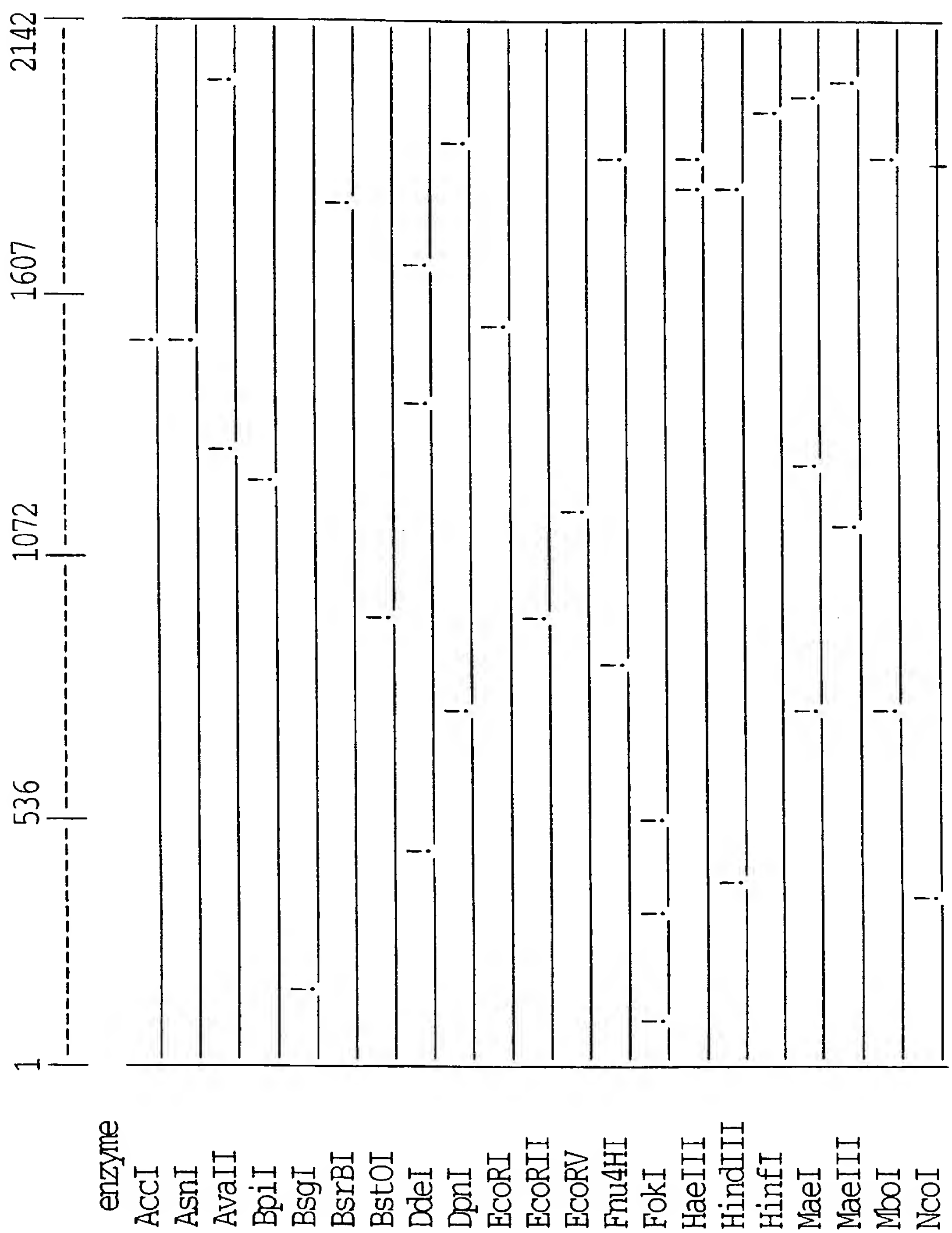
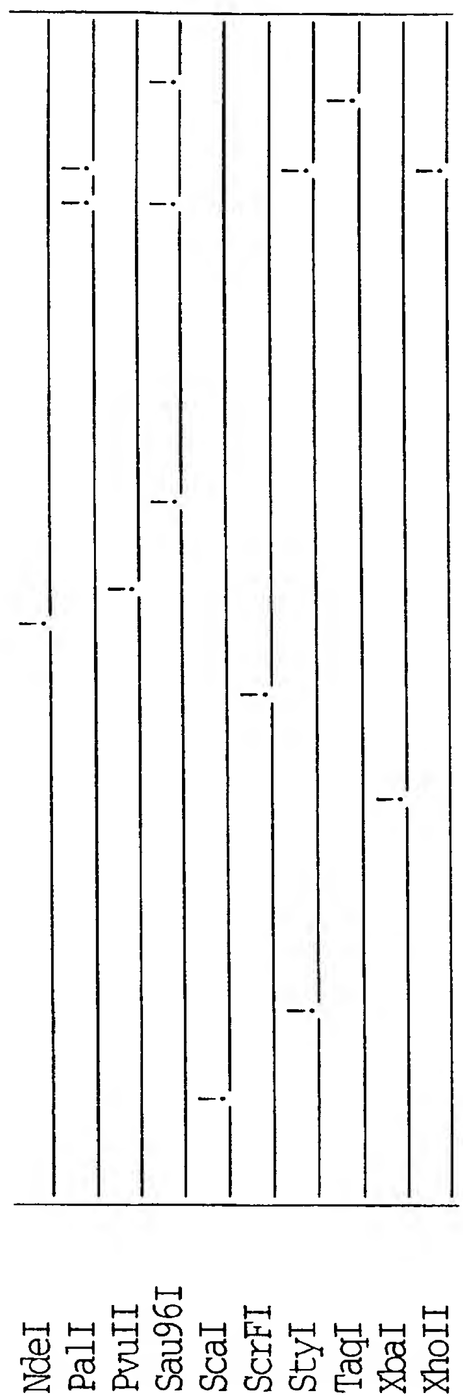


FIG.5B



# FIG.6A

*M. catarrhalis* strain LES1 *tbpB* sequence

```
MET   LYS   HIS   ILE   PRO   LEU   THR   THR   LEU   CYS ...
A T G A A A C A C A T T C C T T T A A C C A C A C T G T G T...
                  10                  20                  30...
                       ... VAL   ALA   ILE   SER   ALA   VAL   LEU   LEU   THR   ALA
                       ... G T G G C A A T C T C T G C C G T C T T A T T A A C C G C T
                       ...                40                  50                  60

CYS   GLY   GLY   SER   GLY   GLY   SER   ASN   PRO   PRO ...
T G T G G T G G C A G T G G T G G T T C A A A T C C A C C T...
                  70                  80                  90...
                       ... ALA   PRO   THR   PRO   ILE   PRO   ASN   ALA   GLY   SER
                       ... G C T C C T A C G C C C A T C C C A A A T G C A G G C A G T
                       ...               100                 110                 120

ALA   GLY   ASN   ALA   GLY   GLY   THR   GLY   ASN   THR ...
G C A G G T A A T G C T G G C G G T A C A G G A A A T A C A...
                 130                 140                 150...
                       ... GLY   GLY   THR   GLY   SER   THR   ASP   ASN   VAL   GLY
                       ... G G C G G T A C T G G C A G T A C T G A T A A T G T A G G C
                       ...               160                 170                 180

ASN   ALA   GLY   GLY   ALA   ASN   SER   GLY   THR   GLY ...
A A T G C T G G C G G T G C A A A C T C T G G T A C A G G C...
                 190                 200                 210...
                       ... ASN   ALA   GLY   ASN   SER   GLY   ASN   ALA   ASN   SER
                       ... A A T G C A G G T A A T T C A G G T A A T G C A A A C T C T
                       ...               220                 230                 240
```

FIG.6B

```
GLY THR GLY SER ALA ASN THR PRO GLU PRO ...
G G T A C A G G C A G T G C C A A C A C A C C A G A A C C A...
        250                 260                 270...
    ... LYS TYR GLN ASP VAL PRO THR ASP LYS ASN
    ... A A A T A T C A A G A T G T G C C A A C C G A T A A A A A T
    ...                 280                 290                 300

GLU LYS GLU GLN VAL SER SER ILE GLN GLU ...
G A A A A A G A A C A A G T T T C A T C C A T T C A A G A A...
        310                 320                 330...
    ... PRO ALA MET GLY TYR ALA MET GLU LEU LYS
    ... C C T G C C A T G G G T T A T G C A A T G G A A T T A A A G
    ...                 340                 350                 360

LEU ARG ASN ALA HIS PRO LEU ASN PRO ASN ...
C T T C G T A A T G C T C A C C C T C T T A A C C C A A A T...
        370                 380                 390...
    ... LYS ASN LYS GLU ALA GLU LYS ARG ILE ALA
    ... A A A A A T A A A G A G G C T G A A A A A C G C A T T G C C
    ...                 400                 410                 420

LEU ASP GLN LYS ASP LEU VAL ALA VAL GLU ...
T T A G A C C A A A A A G A T T T G G T G G C A G T A G A G...
        430                 440                 450...
    ... GLY ASP LEU THR ASN ILE PRO PHE ASP LYS
    ... G G C G A C C T A A C C A A C A T T C C T T T T G A T A A A
    ...                 460                 470                 480
```

FIG.6C

```
ASN LEU ILE GLU TYR LEU LYS LYS SER SER ...
A A T C T T A T T G A A T A C C T T A A A A A A T C A T C C...
          490                     500                     510...
                                 ... GLU VAL VAL SER LYS PHE GLU ALA GLN LYS
                                 ... G A G G T T G T A A G T A A A T T T G A A G C A C A A A A A
                                 ...             520                     530                     540

GLY GLY ILE GLU ASN ASN THR ARG LEU THR ...
G G C G G T A T T G A A A A T A A C A C A A G A C T G A C A...
          550                     560                     570...
                                 ... HIS LYS ASP LEU SER SER GLU GLN LYS GLU
                                 ... C A C A A A G A T T T A T C A T C A G A G C A A A A A G A A
                                 ...             580                     590                     600

ALA LYS VAL LYS GLU ALA LEU ASP ASN ALA ...
G C A A A A G T C A A A G A A G C G T T G G A C A A T G C T...
          610                     620                     630...
                                 ... LEU THR GLN PHE ALA GLN GLU LYS TYR LYS
                                 ... T T A A C T C A A T T T G C C C A A G A A A A A T A C A A G
                                 ...             640                     650                     660

GLU LEU ILE GLU ASN ALA HIS ASP LYS LYS ...
G A G C T A A T T G A G A A C G C C C A T G A T A A A A A A...
          670                     680                     690...
                                 ... SER ASP ALA ARG ASN ARG ASP LEU GLU TYR
                                 ... T C T G A C G C A C G C A A C C G T G A T C T A G A A T A T
                                 ...             700                     710                     720
```

FIG.6D

```
VAL  LYS  SER  GLY  PHE  ASN  TYR  LEU  SER  GLY ...
G T C A A G T C T G G T T T T A A C T A T C T T T C T G G A...
               730                 740                750...
                                   ... TYR  THR  ALA  THR  ASP  HIS  ASP  LYS  LYS  THR
                                   ... T A T A C C G C C A C C G A C C A C G A C A A A A A A A C C
                                   ...             760                770               780

ASN  TYR  ARG  GLY  TYR  TYR  GLY  ALA  LEU  TYR...
A A T T A T C G T G G C T A T T A T G G T G C G T T G T A T...
               790                 800                810...
                                   ... TYR  LYS  GLY  SER  GLU  THR  ALA  LYS  GLU  LEU
                                   ... T A T A A A G G C A G C G A A A C C G C C A A A G A G C T A
                                   ...             820                830               840

PRO  GLN  THR  SER  ALA  LYS  TYR  LYS  GLY  TYR ...
C C A C A A A C A A G T G C A A A A T A T A A A G G T T A T...
               850                 860                870...
                                   ... TRP  ASP  PHE  MET  THR  ASP  ALA  THR  LEU  ASP
                                   ... T G G G A C T T T A T G A C A   G A T G C C A C A C T T G A T
                                   ...             880                890               900

ASN  LYS  TYR  THR  ASP  LEU  PRO  GLY  ILE  ALA ...
A A C A A A T A C A C G G A T T T G C C A G G T A T C G C C...
               910                 920                930...
                                   ... ARG  GLN  THR  GLN  TRP  ARG  SER  LEU  VAL  SER
                                   ... A G A C A A A C C C A G T G G C G T A G T C T T G T T T C T
                                   ...             940                950               960
```

FIG.6E

```
THR   ASP   GLU   TYR   ALA   THR   LEU   LEU   THR   ASP ...
A C T G A T G A G T A T G C A A C G C T C T T G A C A G A C...
         970                     980                     990...
                        ... LYS   ASN   ASN   LYS   PRO   SER   ASP   TYR   ASN   GLY
                        ... A A A A A T A A C A A G C C C A G T G A T T A C A A T G G T
                        ...                1000                  1010                  1020

ALA   TYR   GLY   HIS   SER   SER   GLU   PHE   ASP   VAL ...
G C A T A T G G T C A T A G C A G T G A A T T T G A T G T T...
         1030                    1040                    1050...
                        ... ASN   PHE   ALA   ASP   LYS   LYS   ILE   LYS   GLY   LYS
                        ... A A T T T T G C T G A T A A A A A A T T A A A G G C A A A
                        ...                1060                  1070                  1080

LEU   ILE   SER   ASN   GLN   LEU   SER   GLY   THR   ALA ...
C T T A T C A G T A A T C A G T T A T C A G G C A C A G C T...
         1090                    1100                    1110...
                        ... VAL   THR   ALA   LYS   GLU   ARG   TYR   LYS   ILE   GLU
                        ... G T A A C C G C C A A A G A G C G T T A T A A A A T A G A A
                        ...                1120                  1130                  1140

ALA   ASP   ILE   HIS   GLY   ASN   ARG   PHE   ARG   GLY ...
G C T G A T A T C C A C G G C A A C C G C T T C C G T G G C...
         1150                    1160                    1170...
                        ... SER   ALA   THR   ALA   SER   ASP   LYS   ALA   GLU   ASP
                        ... A G T G C C A C C G C A A G C G A T A A A G C A G A A G A C
                        ...                1180                  1190                  1200
```

FIG.6F

```
SER LYS THR GLN HIS PRO PHE THR SER ASP ...
A G C A A A A C C C A A C A C C C C T T T A C C A G C G A T...
          1210                1220                1230...
    ... ALA THR ASN LYS LEU GLU GLY GLY PHE TYR
    ... G C T A C A A A C A A G C T A G A A G G T G G T T T T T A T
    ...                1240                1250                1260

GLY PRO LYS GLY GLU GLU LEU ALA GLY LYS ...
G G A C C A A A A G G C G A G G A G C T G G C A G G T A A A...
          1270                1280                1290...
    ... PHE LEU THR ASP ASP ASN LYS LEU PHE GLY
    ... T T C T T A A C C G A T G A C A A C A A A C T C T T T G G G
    ...                1300                1310                1320

VAL PHE GLY ALA LYS ARG ASP LYS VAL GLU ...
G T C T T T G G T G C T A A A C G A G A T A A A G T A G A A...
          1330                1340                1350...
    ... LYS THR GLU ALA ILE LEU ASP ALA TYR ALA
    ... A A A A C C G A A G C C A T C T T A G A T G C C T A T G C A
    ...                1360                1370                1380

LEU GLY THR PHE ASN ASN THR ASN LYS ALA ...
C T T G G G A C A T T T A A T A A T A C A A A T A A A G C A...
          1390                1400                1410...
    ... THR THR PHE THR PRO PHE THR LYS LYS GLN
    ... A C C A C A T T C A C C C C A T T T A C C A A A A A A C A A
    ...                1420                1430                1440
```

FIG.6G

```
LEU  ASP  ASN  PHE  GLY  ASN  ALA  LYS  LYS  LEU ...
C T G G A T A A C T T T G G C A A T G C C A A A A A G T T G...
          1450                1460                1470...
                              ... VAL  LEU  GLY  SER  THR  VAL  ILE  ASN  LEU  VAL
                              ... G T C T T G G G T T C T A C C G T C A T T A A T T T G G T G
                              ...           1480                1490                1500

SER  THR  ASP  ALA  THR  LYS  ASN  GLU  PHE  THR ...
T C T A C C G A T G C C A C C A A A A A T G A A T T C A C C...
          1510                1520                1530...
                              ... LYS  LYS  PHE  THR  LYS  ASP  LYS  PRO  THR  SER
                              ... A A A A A A T T C A C C A A A G A C A A G C C A A C T T C T
                              ...           1540                1550                1560

ALA  THR  ASN  LYS  ALA  GLY  GLU  THR  LEU  MET ...
G C C A C A A A C A A A G C G G G C G A G A C T T T G A T G...
          1570                1580                1590...
                              ... VAL  ASN  ASP  GLU  VAL  ILE  VAL  LYS  THR  TYR
                              ... G T G A A T G A T G A A G T T A T C G T C A A A A C C T A T
                              ...           1600                1610                1620

GLY  LYS  ASN  PHE  GLU  TYR  LEU  LYS  PHE  GLY ...
G G C A A A A A C T T T G A A T A C C T A A A A T T T G G T...
          1630                1640                1650...
                              ... GLU  LEU  SER  VAL  GLY  ASP  SER  HIS  SER  VAL
                              ... G A G C T T A G T G T C G G T G A T A G C C A T A G C G T C
                              ...           1660                1670                1680
```

FIG.6H

```
PHE   LEU   GLN   GLY   GLU   ARG   THR   ALA   THR   THR ...
T T T T T A C A A G G C G A A C G C A C C G C T A C C A C A...
            1690                  1700                 1710...
                        ... GLY   GLU   LYS   ALA   VAL   PRO   THR   THR   GLY   LYS
                        ... G G C G A G A A A G C C G T A C C A A C C A C A G G C A A A
                        ...             1720                 1730                 1740

ALA   LYS   TYR   LEU   GLY   ASN   TRP   VAL   GLY   TYR ...
G C C A A A T A T C T G G G G A A C T G G G T A G G A T A C...
            1750                  1760                 1770...
                        ... ILE   THR   GLY   ALA   GLY   THR   GLY   LYS   SER   PHE
                        ... A T C A C A G G A G C G G G C A C A G G A A A A A G C T T T
                        ...             1780                 1790                 1800

ASN   GLU   ALA   GLN   ASP   ILE   ALA   ASP   PHE   ASP ...
A A T G A G G C C C A A G A T A T T G C T G A T T T T G A C...
            1810                  1820                 1830...
                        ... ILE   ASP   PHE   GLU   ARG   LYS   SER   VAL   LYS   GLY
                        ... A T T G A C T T T G A G A G A A A A T C A G T T A A A G G C
                        ...             1840                 1850                 1860

LYS   LEU   THR   THR   GLN   GLY   ARG   THR   ASP   PRO ...
A A A C T G A C C A C C C A A G G C C G C A C A G A T C C T...
            1870                  1880                 1890...
                        ... VAL   PHE   ASN   ILE   LYS   GLY   GLU   ILE   ALA   GLY
                        ... G T C T T T A A C A T C A A A G G T G A A A T T G C A G G C
                        ...             1900                 1910                 1920
```

FIG.6I

```
ASN  GLY  TRP  THR  GLY  LYS  ALA  SER  THR  THR ...
A A T G G C T G G A C A G G C A A A G C C A G C A C C A C C...
          1930                1940                1950...
                    ... LYS  ALA  ASP  ALA  GLY  GLY  TYR  LYS  ILE  ASP
                    ... A A A G C G G A C G C A G G A G G C T A C A A G A T A G A T
                    ...            1960                1970                1980

SER  SER  SER  THR  GLY  LYS  SER  ILE  VAL  ILE ...
T C T A G C A G T A C A G G C A A A T C C A T C G T C A T C...
          1990                2000                2010...
                    ... GLU  ASN  ALA  GLU  VAL  THR  GLY  GLY  PHE  TYR
                    ... G A A A A T G C C G A A G T T A C T G G G G G C T T T T A T
                    ...            2020                2030                2040

GLY  PRO  ASN  ALA  ASN  GLU  MET  GLY  GLY  SER ...
G G T C C A A A T G C A A A C G A G A T G G G C G G G T C A...
          2050                2060                2070...
                    ... PHE  THR  HIS  ASP  THR  ASP  ASP  SER  LYS  ALA
                    ... T T T A C A C A C G A T A C C G A T G A C A G T A A A G C C
                    ...            2080                2090                2100

SER  VAL  VAL  PHE  GLY  THR  LYS  ARG  GLN  GLN ...
T C T G T G G T C T T T G G C A C A A A A A G A C A A C A A...
          2110                2120                2130...
                    ... GLU  VAL  LYS  ***
                    ... G A A G T T A A G T A G
                    ...            2140
```

FIG.7A Alignment of *M. catarrhalis* TbpB protein sequences

```
10        20         30        40        50         60
MKHIPLTTLCVAISAVLLTACGGSGGS-NPPAPTPIPNASGSGNTGN-TGNAGGTDNTAN
...........................-..................-............
...........................-..................-............
...........................-...........GSA..A.G-...T...GS.D.
........................S.GF....S.....GN...A..-A.....GGANS
...........................-...........G.A..A.SG..G..S...A..
                                                       70         80        90       100
                                -AGNTGGT-------------NSGTGSANTPEPKYQDVPTEKNEKDKVSSIQEPAM   4223
                                -.......-------------S......S.S.........TP.N.EQ.........   R1
                                -.......-------------..............K....DE.K..E..G......   M35
                                -V..A..ANSGTGNAGNSGNA...................D....EQ.........   LES1
                                G...A...----------GGA...A...S......K....DE.K.AE..G......   Q8
                                -..S...A-------------S......S.QK........D..K..E..G......   3

     110       120        130        140       150
GYGMALSKINLHNRQDTPLDEKNI-ITLDGKKQVA-EGKKSPLPFSLDVENKLLDGYIA
...........YDQ......A...-..........-DNQ....................
........M...KQ........D.-..........-K.E....................
..A.E.KLR.A.PLNPNKNK.AEKR.A..Q.DL..V..DLTNI..DKNLIEY.KKSSEV
...VE.KLR.WIPQ------.QEEHAKINTNDV.KL..DLKHN..DNSIWQNIKNSKEV
...VE.KLR.WIPQ------.QEEHAKINTNDV.KLEGDLKHN..DNSIWQNIKNSKEV
                                160       170         180        190       200
                                KMNV-ADKNAIGDRIKKG---NKEISDE--ELAKQIKEAVRKSHEFQQV--   4223
                                ...E-.......E...RENEQ..K....--....K...N....P.....--   R1
                                ...E-...........D---..DK.LSKA........D..........--   M35
                                VSKFE.Q.GG.ENNTRLT---H.DL.S.QKE-..-V...---LDNALT.FA   LES1
                                QTVYNQE.QN.EDQ..RE---..QRP.KK.DDVALQAYIEKVLDDRLTELA   Q8
                                QTVYNQE.QN.ENQ...E---...-----LDKTALKALIEKVLDDYLTSLA   3
```

FIG.7B

```
       210       220       230       240        250
LSSLENKIFHSNDGTTKATTRDLKYVDYGYYLANDGNYLTVKTD--KLWNLGPVGGVFY
...IKA.T.....K..................V..A........NP....S........
....K.................Q........V..........--E............
QEKYKEL.ENAH.KKSD.RN...E..KS.FNYLSGYTATDHDK---.TNYR.YY.ALY.
KPIY.KN.NY.H.KQN..R.......RS..IYRSGYSNIIP.----.IAKT.FD.AL..
KPIY.KN.ND.H.KQN..R.......RS..IYRSGYSNIDIQK---.IAKT.FD.AL..
                                   260       270       280       290       300
                                   NGTTTAKELPTQDAVKYKGHWDFMTDVANRRNRFSEVKENSQA      4223
                                   ..S.......................KK.........TY..      R1
                                   ..........................KQ..........L..      M35
                                   K.SE......QTS.-....Y......ATLDNKYTDLPGIAR.T      LES1
                                   Q..Q...Q..VSQ-.....T......AKKGQSFS.FGTSQRL.      Q8
                                   K..Q...Q..VSE-.....T......AKKGQSFS.FERRAGDR      3

       310       320       330       340       350
GWYYGASSKDEYNRLLTKEDSAPDGHSGEYGHSSEFTVNFKEKKLTGKLFSNLQDRH
..W..............A.A...NY...................E.......S.
.R..............D.KNK.ERYN..............D.....E.......SR
Q.-RSLV.T...AT...DKNNK.SDYN.A.......D...AD..IK...I..QLSG-
.DR.S.M.YH..PS...D.KNK..NYN...........D.SK.S.K.E.S..I..G.
--.SAM..H-..PS...DDKNK..NYND..........D.SK.S...G.......H.
                                   360       370       380       390       400
                                   KGNVTKTERYDIDANIHGNRFRGSATASNKNDTSK--HPFTSDAN    4223
                                   .QK....K....K.D.............D.AED..SK.......K    R1
                                   .......K.......Y..........D.AEA..TK.......K      M35
                                   -TA..AK...K.E.D.............D.AED..TQ.......T    LES1
                                   ..S.N..K.......Y..........DTTEA..SK.......K      Q8
                                   ..K....K....N.R............I..DNE..AK.......D    3
```

FIG.7C

```
           410       420       430       440       450
NRLEGGFYGPKGEELAGKFLTNDNKLFGVFGAKRESKAEEKTEAILDAYALGTFNTSN
DK.................D...........Q.GNV-................KPG
.S........NA...........................G.............KNN
.K.................D...........D-.V-.................NT.
.S........NA.........................E.K.............KPG
..........NA.......D...........Q..E.K.-..............KSG
                 460       470       480       490       500
                 -ATT--FTPFTEKQLDNFGNAKKLVLGSTVIDLVPTDATK--NEFTK----    4223
                 -T.NPA..ANSK.E.........................DV...-.----    R1
                 -...--.....K.......................GV..DV.....----    M35
                 K...--.....K.................N..S.....--.....KFTK     LES1
                 -T.NPA..ANSK.E.....................G...DV...-.----    Q8
                 -T.NPA..ANSK.E......IN...........TQG------.D.V.TIDK   3

           510       520       530       540       550
DKPESATNEAGETLMVNDEVSVKTYG--KNFEYLKFGELSIGGSHSVFLQGERTATTG
E..K....K................--...........V................
N..D....K.........K.......YGR..........V.T.N............
...T....K..........I.....--............V.D.............
E..K....K..........I.....YGR........................---
E..ATT..Q...P.T...K.R.QVCC--S.L.H....S....D.N..........K.
                 560       570       580       590       600
                 E--KAVPTTGTAKYLGNWVGYITGKDTGTGT--GKSFTDAQDVADFDI       4223
                 .--.......K............A.SSK.STD..G....K.I.....        R1
                 .--.....K...............S------S...NE.........         M35
                 .--.......K............----A..--....NE...I.....        LES1
                 .--.....E..................S.--....NE...I.....         Q8
                 D.D..M.VA.N...R.T.A..VA.SGNTSKAYEAQQ.A.NANR.E..V       3
```

# FIG.7D

```
     610       620         630       640       650
DFGNKSVSGKLITKGRQDPV--FSITGQIAGNGWTGTASTTKADAGGYKIDSSSTGKS
..EK...N...T..D.....--.N...E........K...AE.N..............
..EK...K...T..D.....--.N...D........K.....................
..ER...K...T.Q..T...--.N.K.E........K.....................
..ER...K...T.Q......--.N................A..NV.............
..A...LT....PNTSS.GKSA.D..AT.D...FS.K.N.PDIET..L....KNSESG

                                      660       670       680                690       700
                                      -IVIKDANVTGGFYGPNANEMGGSFTHNA-------DDSKASVVFGTKRQQEVK-*    4223
                                      -......V.........T........S------GN.G.V.......K.-...K*    R1
                                      -......E..................T-------..............E...-*    M35
                                      -...EN.E..................DT--------.................*    LES1
                                      -...EN.K..................DT-------.............E...-*    Q8
                                      RVIV...I.I......Q...L.....YKSNDAGNQDK..S......ARK.....P*    3
```

# FIG.8A

*M. catarrhalis* strain 4223 *tbpA* - *orf3* - *tbpB* locus gene sequences

```
GATGCCTGCCTTGTGATTGGTTGGGGTGTA...
         10        20        30...
                              ...TCGGTGTATCAAAGTGCAAAAGCCAACAGG
                              ...         40        50        60

          tbpA
          MET ASN GLN SER LYS GLN ASN ...
TGGTCATTGATGAATCAATCAAAACAAAAC...
         70        80        90...
                              ... ASN LYS SER LYS LYS SER LYS GLN VAL LEU
                              ...AACAAATCCAAAAAAATCCAAACAAGTATTA
                              ...         100       110       120

LYS LEU SER ALA LEU SER LEU GLY LEU LEU ...
AAACTTAGTGCCTTGTCTTTGGGTCTGCTT...
         130       140       150...
                              ... ASN ILE THR GLN VAL ALA LEU ALA ASN THR
                              ...AACATCACGCAGGTGGCACTGGCAAACACA
                              ...         160       170       180

THR ALA ASP LYS ALA GLU ALA THR ASP LYS ...
ACGGCCGATAAGGCGGAGGCAACAGATAAG...
         190       200       210...
                              ... THR ASN LEU VAL VAL VAL LEU ASP GLU THR
                              ...ACAAACCTTGTTGTTGTCTTGGATGAAACT
                              ...         220       230       240
```

FIG.8B

```
VAL   VAL   THR   ALA   LYS   LYS   ASN   ALA   ARG   LYS ...
G T T G T A A C A G C G A A G A A A A A C G C C C G T A A A...
            250                     260                  270...
                  ... ALA   ASN   GLU   VAL   THR   GLY   LEU   GLY   LYS   VAL
                  ... G C C A A C G A A G T T A C A G G G C T T G G T A A G G T G
                  ...               280                     290                 300

VAL   LYS   THR   ALA   GLU   THR   ILE   ASN   LYS   GLU ...
G T C A A A A C T G C C G A G A C C A T C A A T A A A G A A...
            310                     320                  330...
                  ... GLN   VAL   LEU   ASN   ILE   ARG   ASP   LEU   THR   ARG
                  ... C A A G T G C T A A A C A T T C G A G A C T T A A C A C G C
                  ...               340                     350                 360

TYR   ASP   PRO   GLY   ILE   ALA   VAL   VAL   GLU   GLN ...
T A T G A C C C T G G C A T T G C T G T G G T T G A G C A A...
            370                     380                  390...
                  ... GLY   ARG   GLY   ALA   SER   SER   GLY   TYR   SER   ILE
                  ... G G T C G T G G G G C A A G C T C A G G C T A T T C T A T T
                  ...               400                     410                 420

ARG   GLY   MET   ASP   LYS   ASN   ARG   VAL   ALA   VAL ...
C G T G G T A T G G A T A A A A A T C G T G T G G C G G T A...
            430                     440                  450...
                  ... LEU   VAL   ASP   GLY   ILE   ASN   GLN   ALA   GLN   HIS
                  ... T T G G T T G A T G G C A T C A A T C A A G C C C A G C A C
                  ...               460                     470                 480
```

FIG.8C

```
TYR  ALA  LEU  GLN  GLY  PRO  VAL  ALA  GLY  LYS ...
T A T G C C C T A C A A G G C C C T G T G G C A G G C A A A...
              490                 500                510...
                                 ... ASN  TYR  ALA  ALA  GLY  GLY  ALA  ILE  ASN  GLU
                                 ... A A T T A T G C C G C A G G T G G G G C A A T C A A C G A A
                                 ...             520              530              540

ILE  GLU  TYR  GLU  ASN  VAL  ARG  SER  VAL  GLU ...
A T A G A A T A C G A A A A T G T C C G C T C C G T T G A G...
              550                 560                570...
                                 ... ILE  SER  LYS  GLY  ALA  ASN  SER  SER  GLU  TYR
                                 ... A T T A G T A A A G G T G C A A A T T C A A G T G A A T A C
                                 ...             580              590              600

GLY  SER  GLY  ALA  LEU  SER  GLY  SER  VAL  ALA ...
G G C T C T G G G G C A T T A T C T G G C T C T G T G G C A...
              610                 620                630...
                                 ... PHE  VAL  THR  LYS  THR  ALA  ASP  ASP  ILE  ILE
                                 ... T T T G T T A C C A A A A C C G C C G A T G A C A T C A T C
                                 ...             640              650              660

LYS  ASP  GLY  LYS  ASP  TRP  GLY  VAL  GLN  THR ...
A A A G A T G G T A A A G A T T G G G G C G T G C A G A C C...
              670                 680                690...
                                 ... LYS  THR  ALA  TYR  ALA  SER  LYS  ASN  ASN  ALA
                                 ... A A A A C C G C C T A T G C C A G T A A A A A T A A C G C A
                                 ...             700              710              720
```

FIG.8D

```
TRP   VAL   ASN   SER   VAL   ALA   ALA   ALA   GLY   LYS ...
T G G G T T A A T T C T G T G G C A G C A G C A G G C A A G...
              730                     740                750...
                            ... ALA   GLY   SER   PHE   SER   GLY   LEU   ILE   ILE   TYR
                            ... G C A G G T T C T T T T A G C G G T C T T A T C A T C T A C
                            ...               760                 770                 780

THR   ASP   ARG   ARG   GLY   GLN   GLU   TYR   LYS   ALA ...
A C C G A C C G C C G T G G T C A A G A A T A C A A G G C A...
              790                     800                810...
                            ... HIS   ASP   ASP   ALA   TYR   GLN   GLY   SER   GLN   SER
                            ... C A T G A T G A T G C C T A T C A G G G T A G C C A A A G T
                            ...               820                 830                 840

PHE   ASP   ARG   ALA   VAL   ALA   THR   THR   ASP   PRO ...
T T T G A T A G A G C G G T G G C A A C C A C T G A C C C A...
              850                     860                870...
                            ... ASN   ASN   ARG   THR   PHE   LEU   ILE   ALA   ASN   GLU
                            ... A A T A A C C G A A C A T T T T T A A T A G C A A A T G A A
                            ...               880                 890                 900

CYS   ALA   ASN   GLY   ASN   TYR   GLU   ALA   CYS   ALA ...
T G T G C C A A T G G T A A T T A T G A G G C G T G T G C T...
              910                     920                930...
                            ... ALA   GLY   GLY   GLN   THR   LYS   LEU   GLN   ALA   LYS
                            ... G C T G G C G G T C A A A C C A A A C T T C A A G C C A A G
                            ...               940                 950                 960
```

FIG.8E

```
PRO  THR  ASN  VAL  ARG  ASP  LYS  VAL  ASN  VAL ...
C C A A C C A A T G T G C G T G A T A A G G T C A A T G T C...
          970                 980                 990...
                         ... LYS  ASP  TYR  THR  GLY  PRO  ASN  ARG  LEU  ILE
                         ... A A A G A T T A T A C A G G T C C T A A C C G C C T T A T C
                         ...           1000                1010                1020

PRO  ASN  PRO  LEU  THR  GLN  ASP  SER  LYS  SER ...
C C A A A C C C A C T C A C C C A A G A C A G C A A A T C C...
          1030                1040                1050...
                         ... LEU  LEU  LEU  ARG  PRO  GLY  TYR  GLN  LEU  ASN
                         ... T T A C T G C T T C G C C C A G G T T A T C A G C T A A A C
                         ...           1060                1070                1080

ASP  LYS  HIS  TYR  VAL  GLY  GLY  VAL  TYR  GLU ...
G A T A A G C A C T A T G T C G G T G G T G T G T A T G A A...
          1090                1100                1110...
                         ... ILE  THR  LYS  GLN  ASN  TYR  ALA  MET  GLN  ASP
                         ... A T C A C C A A A C A A A A C T A C G C C A T G C A A G A T
                         ...           1120                1130                1140

LYS  THR  VAL  PRO  ALA  TYR  LEU  ALA  VAL  HIS ...
A A A A C C G T G C C T G C T T A T C T G G C G G T T C A T...
          1150                1160                1170...
                         ... ASP  ILE  GLU  LYS  SER  ARG  LEU  SER  ASN  HIS
                         ... G A C A T T G A A A A A T C A A G G C T C A G C A A C C A T
                         ...           1180                1190                1200
```

FIG.8F

```
ALA   GLN   ALA   ASN   GLY   TYR   TYR   GLN   GLY   ASN ...
G C C C A A G C C A A T G G C T A T T A T C A A G G C A A T...
              1210                1220                1230...
                              ... ASN   LEU   GLY   GLU   ARG   ILE   ARG   ASP   THR   ILE
                              ... A A T C T T G G T G A A C G C A T T C G T G A T A C C A T T
                              ...              1240                1250                1260

GLY   PRO   ASP   SER   GLY   TYR   GLY   ILE   ASN   TYR ...
G G G C C A G A T T C A G G T T A T G G C A T C A A C T A T...
              1270                1280                1290...
                              ... ALA   HIS   GLY   VAL   PHE   TYR   ASP   GLU   LYS   HIS
                              ... G C T C A T G G C G T A T T T T A T G A T G A A A A A C A C
                              ...              1300                1310                1320

GLN   LYS   ASP   ARG   LEU   GLY   LEU   GLU   TYR   VAL ...
C A A A A A G A C C G C C T A G G G C T T G A A T A T G T T...
              1330                1340                1350...
                              ... TYR   ASP   SER   LYS   GLY   GLU   ASN   LYS   TRP   PHE
                              ... T A T G A C A G C A A A G G T G A A A A T A A A T G G T T T
                              ...              1360                1370                1380

ASP   ASP   VAL   ARG   VAL   SER   TYR   ASP   LYS   GLN ...
G A T G A T G T G C G T G T G T C T T A T G A T A A G C A A...
              1390                1400                1410...
                              ... ASP   ILE   THR   LEU   ARG   SER   GLN   LEU   THR   ASN
                              ... G A C A T T A C G C T A C G C A G C C A G C T G A C C A A C
                              ...              1420                1430                1440
```

FIG.8G

```
THR HIS CYS SER THR TYR PRO HIS ILE ASP ...
A C G C A C T G T T C A A C C T A T C C G C A C A T T G A C...
          1450                          1460                1470...
                                 ... LYS ASN CYS THR PRO ASP VAL ASN LYS PRO
                                 ...A A A A A T T G T A C G C C T G A T G T C A A T A A A C C T
                                 ...              1480                1490                1500

PHE SER VAL LYS GLU VAL ASP ASN ASN ALA ...
T T T T C G G T A A A A G A G G T G G A T A A C A A T G C C...
          1510                          1520                1530...
                                 ... TYR LYS GLU GLN HIS ASN LEU ILE LYS ALA
                                 ... T A C A A A G A A C A G C A C A A T T T A A T C A A A G C C
                                 ...              1540                1550                1560

VAL PHE ASN LYS LYS MET ALA LEU GLY SER ...
G T C T T T A A C A A A A A A A T G G C G T T G G G C A G T...
          1570                          1580                1590...
                                 ... THR HIS HIS HIS ILE ASN LEU GLN VAL GLY
                                 ... A C G C A T C A T C A C A T C A A C C T G C A A G T T G G C
                                 ...              1600                1610                1620

TYR ASP LYS PHE ASN SER SER LEU SER ARG ...
T A T G A T A A A T T C A A T T C A A G C C T G A G C C G T...
          1630                          1640                1650...
                                 ... VAL GLU TYR ARG LEU ALA THR HIS GLN SER
                                 ... G T A G A A T A T C G T T T G G C A A C C C A T C A G T C T
                                 ...              1660                1670                1680
```

FIG.8H

```
TYR GLN LYS LEU ASP TYR THR PRO PRO SER ...
TATCAAAAACTTGATTACACCCCACCAAGT...
        1690      1700      1710...
... ASN PRO LEU PRO ASP LYS PHE LYS PRO ILE
... AACCCTTTGCCAGATAAGTTTAAGCCCATT
...       1720      1730      1740

LEU GLY SER ASN ASN LYS PRO ILE CYS LEU ...
TTAGGTTCAAACAACAAACCCATTTGCCTT...
        1750      1760      1770 ...
... ASP ALA TYR GLY TYR GLY HIS ASP HIS PRO
... GATGCTTATGGTTATGGTCATGACCATCCA
...       1780      1790      1800

GLN ALA CYS ASN ALA LYS ASN SER THR TYR ...
CAGGCTTGTAACGCCAAAAACAGCACTTAT...
        1810      1820      1830...
... GLN ASN PHE ALA ILE LYS LYS GLY ILE GLU
... CAAAATTTTGCCATCAAAAAAGGCATAGAG
...       1840      1850      1860

GLN TYR ASN GLN LYS THR ASN THR ASP LYS ...
CAATACAACCAAAAAACCAATACCGATAAG...
        1870      1880      1890...
... ILE ASP TYR GLN ALA ILE ILE ASP GLN TYR
... ATTGATTATCAAGCCATCATTGACCAATAT
...       1900      1910      1920
```

FIG.8I

```
ASP   LYS   GLN   ASN   PRO   ASN   SER   THR   LEU   LYS ...
G A T A A A C A A A A C C C C A A C A G C A C C C T A A A A...
              1930                1940                1950...
                                  ... PRO   PHE   GLU   LYS   ILE   LYS   GLN   SER   LEU   GLY
                                  ... C C C T T T G A G A A A A T C A A A C A A A G T T T G G G G
                                  ...               1960                1970                1980

GLN   GLU   LYS   TYR   ASN   LYS   ILE   ASP   GLU   LEU ...
C A A G A A A A A T A C A A C A A G A T A G A C G A A C T T...
              1990                2000                2010...
                                  ... GLY   PHE   LYS   ALA   TYR   LYS   ASP   LEU   ARG   ASN
                                  ... G G C T T T A A A G C T T A T A A A G A T T T A C G C A A C
                                  ...               2020                2030                2040

GLU   TRP   ALA   GLY   TRP   THR   ASN   ASP   ASN   SER ...
G A A T G G G C G G G T T G G A C T A A T G A C A A C A G C...
              2050                2060                2070...
                                  ... GLN   GLN   ASN   ALA   ASN   LYS   GLY   THR   ASP   ASN
                                  ... C A A C A A A A T G C C A A T A A A G G C A C G G A T A A T
                                  ...               2080                2090                2100

ILE   TYR   GLN   PRO   ASN   GLN   ALA   THR   VAL   VAL ...
A T C T A T C A G C C A A A T C A A G C A A C T G T G G T C...
              2110                2120                2130...
                                  ... LYS   ASP   ASP   LYS   CYS   LYS   TYR   SER   GLU   THR
                                  ... A A A G A T G A C A A A T G T A A A T A T A G C G A G A C C
                                  ...               2140                2150                2160
```

FIG.8J

```
ASN   SER   TYR   ALA   ASP   CYS   SER   THR   THR   ARG ...
A A C A G C T A T G C T G A T T G C T C A A C C A C T C G C...
          2170                2180                2190...
                        ... HIS   ILE   SER   GLY   ASP   ASN   TYR   PHE   ILE   ALA
                        ... C A C A T C A G T G G T G A T A A T T A T T T C A T C G C T
                        ...           2200                2210                2220

LEU   LYS   ASP   ASN   MET   THR   ILE   ASN   LYS   TYR ...
T T A A A A G A C A A C A T G A C C A T C A A T A A A T A T...
          2230                2240                2250...
                        ... VAL   ASP   LEU   GLY   LEU   GLY   ALA   ARG   TYR   ASP
                        ... G T T G A T T T G G G G C T G G G T G C T C G C T A T G A C
                        ...           2260                2270                2280

ARG   ILE   LYS   HIS   LYS   SER   ASP   VAL   PRO   LEU ...
A G A A T C A A A C A C A A A T C T G A T G T G C C T T T G...
          2290                2300                2310...
                        ... VAL   ASP   ASN   SER   ALA   SER   ASN   GLN   LEU   SER
                        ... G T A G A C A A C A G T G C C A G C A A C C A G C T G T C T
                        ...           2320                2330                2340

TRP   ASN   PHE   GLY   VAL   VAL   VAL   LYS   PRO   THR ...
T G G A A T T T T G G C G T G G T C G T C A A G C C C A C C...
          2350                2360                2370...
                        ... ASN   TRP   LEU   ASP   ILE   ALA   TYR   ARG   SER   SER
                        ... A A T T G G C T G G A C A T C G C T T A T A G A A G C T C G
                        ...           2380                2390                2400
```

FIG.8K

```
GLN GLY PHE ARG MET PRO SER PHE SER GLU ...
C A A G G C T T T C G C A T G C C A A G T T T T T C T G A A...
        2410                2420                2430...
    ... MET TYR GLY GLU ARG PHE GLY VAL THR ILE
    ... A T G T A T G G C G A A C G C T T T G G C G T A A C C A T C
    ...                2440                2450                2460

GLY LYS GLY THR GLN HIS GLY CYS LYS GLY ...
G G T A A A G G C A C G C A A C A T G G C T G T A A G G G T...
        2470                2480                2490...
    ... LEU TYR TYR ILE CYS GLN GLN THR VAL HIS
    ... C T T T A T T A C A T T T G T C A G C A G A C T G T C C A T
    ...                2500                2510                2520

GLN THR LYS LEU LYS PRO GLU LYS SER PHE ...
C A A A C C A A G C T A A A A C C T G A A A A A T C C T T T...
        2530                2540                2550...
    ... ASN GLN GLU ILE GLY ALA THR LEU HIS ASN
    ... A A C C A A G A A A T C G G A G C G A C T T T A C A T A A C
    ...                2560                2570                2580

HIS LEU GLY SER LEU GLU VAL SER TYR PHE ...
C A C T T A G G C A G T C T T G A G G T T A G T T A T T T T...
        2590                2600                2610...
    ... LYS ASN ARG TYR THR ASP LEU ILE VAL GLY
    ... A A A A A T C G C T A T A C C G A T T T G A T T G T T G G T
    ...                2620                2630                2640
```

FIG.8L

```
LYS SER GLU GLU ILE ARG THR LEU THR GLN ...
A A A A G T G A A G A G A T T A G A A C C C T A A C C C A A...
2650 2660 2670...
... GLY ASP ASN ALA GLY LYS GLN ARG GLY LYS
... G G T G A T A A T G C A G G C A A A C A G C G T G G T A A A
... 2680 2690 2700

GLY ASP LEU GLY PHE HIS ASN GLY GLN ASP ...
G G T G A T T T G G G C T T T C A T A A T G G A C A A G A T...
2710 2720 2730 ...
... ALA ASP LEU THR GLY ILE ASN ILE LEU GLY
... G C T G A T T T G A C A G G A A T T A A C A T T C T T G G C
... 2740 2750 2760

ARG LEU ASP LEU ASN ALA ALA ASN SER ARG ...
A G A C T T G A C C T A A A C G C T G C C A A T A G T C G C...
2770 2780 2790...
... LEU PRO TYR GLY LEU TYR SER THR LEU ALA
... C T T C C C T A T G G A T T A T A C T C A A C A C T G G C T
... 2800 2810 2820

TYR ASN LYS VAL ASP VAL LYS GLY LYS THR ...
T A T A A C A A A G T T G A T G T T A A A G G A A A A A C C...
2830 2840 2850...
... LEU ASN PRO THR LEU ALA GLY THR ASN ILE
... T T A A A C C C A A C T T T G G C A G G A A C A A A C A T A
... 2860 2870 2880
```

FIG.8M

```
LEU   PHE   ASP   ALA   ILE   GLN   PRO   SER   ARG   TYR ...
C T G T T T G A T G C C A T C C A G C C A T C T C G T T A T...
              2890                2900                2910...
                              ... VAL   VAL   GLY   LEU   GLY   TYR   ASP   ALA   PRO   SER
                              ... G T G G T G G G C T T G G C T A T G A T G C C C C A A G C
                              ...             2920                2930                2940

GLN   LYS   TRP   GLY   ALA   ASN   ALA   ILE   PHE   THR ...
C A A A A A T G G G G A G C A A A C G C C A T A T T T A C C...
              2950                2960                2970...
                              ... HIS   SER   ASP   ALA   LYS   ASN   PRO   SER   GLU   LEU
                              ... C A T T C T G A T G C C A A A A A T C C A A G C G A G C T T
                              ...             2980                2990                3000

LEU   ALA   ASP   LYS   ASN   LEU   GLY   ASN   GLY   ASN ...
T T G G C A G A T A A G A A C T T A G G T A A T G G C A A C...
              3010                3020                3030...
                              ... ILE   GLN   THR   LYS   GLN   ALA   THR   LYS   ALA   LYS
                              ... A T T C A A A C A A A A C A A G C C A C C A A A G C A A A A
                              ...             3040                3050                3060

SER   THR   PRO   TRP   GLN   THR   LEU   ASP   LEU   SER ...
T C C A C G C C G T G G C A A A C A C T T G A T T T G T C A...
              3070                3080                3090...
                              ... GLY   TYR   VAL   ASN   ILE   LYS   ASP   ASN   PHE   THR
                              ... G G T T A T G T A A A C A T A A A A G A T A A T T T T A C C
                              ...             3100                3110                3120
```

FIG.8N

```
LEU ARG ALA GLY VAL TYR ASN VAL PHE ASN ...
T T G C G T G C T G G C G T G T A C A A T G T A T T T A A T...
            3130                3140                3150...
                              ... THR TYR TYR THR THR TRP GLU ALA LEU ARG
                              ... A C C T A T T A C A C C A C T T G G G A G G C T T T A C G C
                              ...             3160                3170                3180

GLN THR ALA LYS GLY ALA VAL ASN GLN HIS ...
C A A A C A G C A A A A G G G G C G G T C A A T C A G C A T...
            3190                3200                3210...
                              ... THR GLY LEU SER GLN ASP LYS HIS TYR GLY
                              ... A C A G G A C T G A G C C A A G A T A A G C A T T A T G G T
                              ...             3220                3230                3240

ARG TYR ALA ALA PRO GLY ARG ASN TYR GLN ...
C G C T A T G C C G C T C C T G G A C G C A A T T A C C A A...
            3250                3260                3270...
                              ... LEU ALA LEU GLU MET LYS PHE ***
                              ... T T G G C A C T T G A A A T G A A G T T T T A A C C A G T G
                              ...             3280                3290                3300

G C T T T G A T G T G A T T T T G G C A T G C C A A A T C C...
            3310                3320                3330...

                              ...C A A T C A A C C A A T G A A T A A A G C C C C C A T T A C
                              ...             3340                3350                3360
```

FIG.8O

```
CATGAGGGCTTTATTTTATCATCGCTGAGT...
        3370      3380      3390...

                    ...ATGCTCTTAGCGGTCATCACTCAGATTAGT
                    ...     3400      3410      3420

CATTAATTTATTAGCGATTAATTTATTAGT...
        3430      3440      3450...

                    ...AATCACGCTGCTCTTTGATGATTTTAAGTG
                    ...     3460      3470      3480

ATGGGTATTCAAGAACGATGTCATACTCAG...
        3490      3500      3510...

                    ...CACCGTTTTTATAGGCTTCTACTTCAAAGA
                    ...     3520      3530      3540

CAGGCTTGCCTAAAAAGTCATCAACTTCTA...
        3550      3560      3570...

                    ...TATCGCCGACTTGATAGCCACGAGCAGCAA
                    ...     3580      3590      3600

GCATTTGAATGGCTTTTTGACGATTTTGGG...
        3610      3620      3630...

                    ...CAAAGTTGCTGTCGCCATAAGCTTGTGCTT
                    ...     3640      3650      3660
```

FIG.8P

```
T A A T A C G G T C G T T A G C A A C T G C G G T G G T A G...
          3670                3680                3690...

                       ... A G A T A C C A A C G G C A G G C A A C A A A A C A G C A G
                       ...                3700                3710                3720

C A C T T A G T A C G C C A G C C A A C A G T T T A T T G G...
          3730                3740                3750...

                       ... T T A A A T T T T T C A T A G T A G T T T C C T A A T T A T
                       ...                3760                3770                3780

T A T C A T T G T A A T T C A T G T T T A T C G T T A T A A...
          3790                3800                3810...

                       ... A C A A T C G T T A T A A A T A A C T G T G T C G T G A T A
                       ...                3820                3830                3840

A C C A T T A A T C A C A A G T G G G T T A A A T G C C T T...
          3850                3860                3870...

                       ... T T G C C C A A T G G C A A A T A G G C A C A A T G C T C T
                       ...                3880                3890                3900

G C T T G T T C T A T G A T G G T C T A T T A T G A T C A T...
          3910                3920                3930...

                       ... C A T T T T A T T G A C C T A T T T T T T T A A T C G T A A
                       ...                3940                3950                3960
```

FIG.8Q

```
T G T T T G T T T G A T G T T A G T A T A A A T T T T A T C...
          3970                3980                3990...

                  ... A A T C A A A C A A T C A C A A A T T A T A T C A A T C A T
                  ...          4000                4010                4020

A G A C G G T A A A C A G G C T T C A T A T T T T A C G C A...
          4030                4040                4050...

                  ... T A T T T C C C C A G A T G T C T G T A G T G T T T C A T A
                  ...          4060                4070                4080

G A T G A T T T G T A A A A C A A T T G T C G G T C A T T A...
          4090                4100                4110...

                  ... T T A T C A A T T G T A A A C T G A T G G C T A A T T T G T
                  ...          4120                4130                4140

A A C C T T A T G G C T A A T G A T A A T A T G A A T A A A...
          4150                4160                4170...

                  ... G C G T T A T A C T G T A T C A A A G A A T G A G T A A A A
                  ...          4180                4190                4200

A C C A T C A A T G G T A T C T T A T T T A T C A T C A G G...
          4210                4220                4230...

                  ... T T G T G T T A A T A A G A T G C C A A T T A A G C G A C T
                  ...          4240                4250                4260
```

FIG.8R

```
AATTTTGTAAATTAATTAATAATCATTCAT...
        4270      4280      4290...

                      ...ATTTGTATTTTTAAATACCATAAAAAATGG
                      ...        4300      4310      4320

          orf3
          MET LEU ALA PHE LEU ILE GLY ALA ...
TAAAATATGCTCGCTTTTTTGATAGGAGCT...
        4330      4340      4350...

                      ... VAL MET THR ILE THR PRO VAL TYR THR THR
                      ...GTCATGACAATCACGCCTGTTTATACCACA
                      ...        4360      4370      4380

PHE THR PRO THR LYS THR PRO ILE LYS PHE ...
TTCACCCCCACCAAAACACCCATAAAATTT...
        4390      4400      4410...
                      ... PHE MET ALA GLY LEU THR PHE LEU ILE ALA
                      ...TTTATGGCTGGCTTGACTTTTCTAATCGCT
                      ...        4420      4430      4440

HIS ILE SER HIS ALA ASP ASP GLY ARG THR ...
CATATCAGCCATGCCGATGATGGTCGCACC...
        4450      4460      4470...
                      ... ASP ASN GLN GLU LEU ILE ASN GLN GLU ILE
                      ...GACAATCAAGAGCTAATCAATCAAGAAATA
                      ...        4480      4490      4500
```

FIG.8S

```
ALA THR LEU GLU PRO ILE ILE ASN HIS ALA ...
G C C A C C C T T G A A C C C A T C A T T A A C C A T G C T...
          4510                4520                4530...
                          ... GLN PRO GLU LEU LEU SER HIS ASP ALA LEU
                          ... C A G C C T G A G T T A T T G T C C C A T G A T G C A T T A
                          ...           4540                4550                4560

THR PRO LYS ILE GLU PRO ILE LEU ALA GLN ...
A C A C C A A A A A T A G A A C C A A T A C T G G C A C A A...
          4570                4580                4590...
                          ... THR PRO ASN PRO ALA GLU ASP THR LEU ILE
                          ... A C A C C A A A T C C T G C C G A A G A T A C G C T C A T C
                          ...           4600                4610                4620

ALA ASP GLU ALA LEU LEU LEU ASP ASN PRO ...
G C C G A T G A G G C G T T A C T G C T T G A T A A C C C T...
          4630                4640                4650...
                          ... ASP LEU LEU ASN HIS ALA LEU ASN SER ALA
                          ... G A T T T G C T C A A T C A C G C C C T A A A T T C T G C T
                          ...           4660                4670                4680

VAL MET THR ASN HIS MET ALA GLY VAL HIS ...
G T C A T G A C C A A T C A T A T G G C A G G C G T T C A C...
          4690                4700                4710...
                          ... ALA LEU LEU PRO ILE TYR GLN LYS LEU PRO
                          ... G C A T T A T T G C C C A T T T A T C A A A A A C T G C C C
                          ...           4720                4730                4740
```

FIG.8T

```
LYS   ASP   HIS   GLN   ASN   GLY   ILE   LEU   LEU   GLY ...
A A A G A C C A T C A A A A T G G C A T T T T A C T T G G G...
            4750                4760                4770...
                  ... TYR   ALA   ASN   ALA   LEU   ALA   ALA   LEU   ASP   LYS
                  ... T A T G C C A A T G C C T T G G C T G C T T T G G A T A A G
                  ...             4780                4790                4800

GLY   ASN   ALA   LYS   LYS   ALA   ILE   ASP   GLU   LEU ...
G G C A A C G C C A A A A A A G C C A T T G A T G A G C T A...
            4810                4820                4830...
                  ... ARG   ARG   ILE   ILE   ALA   ILE   MET   PRO   GLU   TYR
                  ... C G T C G C A T C A T C G C C A T C A T G C C T G A A T A T
                  ...             4840                4850                4860

ASN   VAL   VAL   ARG   PHE   HIS   LEU   ALA   ARG   ALA ...
A A T G T G G T G C G T T T T C A T C T G G C A A G G G C A...
            4870                4880                4890...
                  ... LEU   PHE   MET   ASP   LYS   GLN   ASN   GLU   ALA   ALA
                  ... T T A T T T A T G G A C A A A C A A A A T G A A G C C G C C
                  ...             4900                4910                4920

LEU   ASP   GLN   PHE   ASN   LYS   LEU   HIS   ALA   ASP ...
C T T G A C C A G T T T A A T A A A T T A C A T G C T G A C...
            4930                4940                4950...
                  ... ASN   LEU   PRO   GLU   GLU   VAL   ARG   GLN   VAL   VAL
                  ... A A C T T G C C A G A G G A G G T G C G G C A G G T T G T T
                  ...             4960                4970                4980
```

FIG.8U

```
GLY   GLN   TYR   ARG   GLN   ALA   LEU   LYS   GLN   ARG ...
G G G C A G T A C A G A C A A G C G C T A A A A C A A C G A...
          4990                5000                5010...
                  ... ASP   SER   TRP   THR   TRP   GLN   VAL   GLY   MET   ASN
                  ... G A T T C A T G G A C A T G G C A A G T A G G C A T G A A T
                  ...             5020                5030                5040

LEU   ALA   LYS   GLU   ASP   ASN   ILE   ASN   GLN   THR ...
C T G G C C A A A G A A G A C A A C A T C A A T C A A A C C...
          5050                5060                5070...
                  ... PRO   LYS   ASN   THR   THR   GLN   GLY   GLN   TRP   THR
                  ... C C C A A A A A C A C C A C G C A A G G T C A A T G G A C T
                  ...                 5080                5090                5100

PHE   ASP   LYS   PRO   ILE   ASP   ALA   ILE   THR   LEU ...
T T T G A C A A A C C C A T T G A C G C C A T C A C C C T A...
          5110                5120                5130...
                  ... SER   TYR   GLN   LEU   GLY   ALA   ASP   LYS   LYS   TRP
                  ... A G C T A C C A A T T G G G G C G G A T A A A A A G T G G
                  ...                 5140                5150                5160

SER   LEU   PRO   LYS   GLY   ALA   TYR   VAL   GLY   ALA ...
T C T T T G C C C A A A G G G G C A T A T G T G G G A G C G...
          5170                5180                5190...
                  ... ASN   ALA   GLN   ILE   TYR   GLY   LYS   HIS   HIS   GLN
                  ... A A C G C C C A A A T C T A T G G C A A A C A T C A T C A A
                  ...                 5200                5210                5220
```

FIG.8V

```
ASN   HIS   LYS   LYS   TYR   ASN   ASP   HIS   TRP   GLY ...
A A T C A C A A A A A A T A C A A C G A C C A T T G G G G C...
            5230                          5240                      5250...
                    ... ARG   LEU   GLY   ALA   ASN   LEU   GLY   PHE   ALA   ASP
                    ... A G A C T G G G G G C A A A T T T G G G C T T T G C T G A T
                    ...              5260                5270                5280

ALA   LYS   LYS   ASP   LEU   SER   ILE   GLU   THR   TYR ...
G C C A A A A A A G A C C T T A G C A T T G A G A C C T A T...
            5290                          5300                      5310...
                    ... GLY   GLU   LYS   ARG   PHE   TYR   GLY   HIS   GLU   ARG
                    ... G G T G A A A A A A G A T T T T A T G G G C A T G A G C G T
                    ...              5320                5330                5340

TYR   THR   ASP   THR   ILE   GLY   ILE   ARG   MET   SER ...
T A T A C C G A C A C C A T T G G C A T A C G C A T G T C G...
            5350                          5360                      5370...
                    ... VAL   ASP   TYR   ARG   ILE   ASN   PRO   LYS   PHE   GLN
                    ... G T T G A T T A T A G A A T C A A C C C A A A A T T T C A A
                    ...              5380                5390                5400

SER   LEU   ASN   ALA   ILE   ASP   ILE   SER   ARG   LEU ...
A G C C T A A A C G C C A T A G A C A T A T C A C G C C T A...
            5410                          5420                      5430...
                    ... THR   ASN   HIS   ARG   THR   PRO   ARG   ALA   ASP   SER
                    ... A C C A A C C A T C G G A C G C C T A G G G C T G A C A G T
                    ...              5440                5450                5460
```

FIG.8W

```
ASN ASN THR LEU TYR SER THR SER LEU ILE ...
A A T A A C A C T T T A T A C A G T A C C T C A T T G A T T...
          5470                          5480                5490...
                         ... TYR TYR PRO ASN ALA THR ARG TYR TYR LEU
                         ... T A T T A C C C A A A T G C C A C A C G C T A T T A T C T T
                         ...         5500                5510                5520

LEU GLY ALA ASP PHE TYR ASP GLU LYS VAL ...
T T G G G G G C A G A C T T T T A T G A T G A A A A A G T G...
          5530                          5540                5550...
                         ... PRO GLN ASP PRO SER ASP SER TYR GLN ARG
                         ... C C A C A A G A C C C A T C T G A C A G T T A T C A A C G C
                         ...         5560                5570                5580

ARG GLY ILE ARG THR ALA TRP GLY GLN GLU ...
C G T G G C A T A C G C A C A G C G T G G G G G C A A G A A...
          5590                          5600                5610...
                         ... TRP ALA GLY GLY LEU SER SER ARG ALA GLN
                         ... T G G G C G G G T G G T C T T T C A A G C C G T G C C C A A
                         ...         5620                5630                5640

ILE SER ILE ASN LYS ARG HIS TYR GLN GLY ...
A T C A G C A T C A A C A A A C G C C A T T A C C A A G G G...
          5650                          5660                5670...
                         ... ALA ASN LEU THR SER GLY GLY GLN ILE ARG
                         ... G C A A A C C T A A C C A G C G G T G G A C A A A T T C G C
                         ...         5680                5690                5700
```

FIG.8X

```
HIS ASP LYS GLN MET GLN ALA SER LEU SER ...
C A T G A T A A A C A G A T G C A A G C G T C T T T A T C G...
5710 5720 5730...
... LEU TRP HIS ARG ASP ILE HIS LYS TRP GLY
... C T T T G G C A C A G A G A C A T T C A C A A A T G G G G C
... 5740 5750 5760

ILE THR PRO ARG LEU THR ILE SER THR ASN ...
A T C A C G C C A C G G C T G A C C A T C A G C A C A A A C...
5770 5780 5790...
... ILE ASN LYS SER ASN ASP ILE LYS ALA ASN
... A T C A A T A A A A G C A A T G A C A T C A A G G C A A A T
... 5800 5810 5820

TYR HIS LYS ASN GLN MET PHE VAL GLU PHE ...
T A T C A C A A A A A T C A A A T G T T T G T T G A G T T T...
5830 5840 5850...
... SER ARG ILE PHE ***
... A G T C G C A T T T T T T T G A T G G G A T A A G C A C G C C
... 5860 5870 5880

C T A C T T T T G T T T T T G T A A A A A A A T G T G C C A...
5890 5900 5910...

... T C A T A G A C A A T A T C A A G A A A A A A T C A A G A A
... 5920 5930 5940
```

FIG.8Y

```
A A A A A G A T T A C A A A T T T A A T G A T A A T T G T T...
          5950                5960                5970...

                    ... A T T G T T T A T G T T A T T A T T T A T C A A T G T A A A
                    ...               5980                5990                6000

T T T G C C G T A T T T T G T C T A T C A T A A A T G C A T...
          6010                6020                6030...

                    ... T T A T C A A A T G C T C A A A T A A A T A C G C C A A A T
                    ...               6040                6050                6060

G C A C A T T G T C A G C A T G C C A A A A T A G G C A T C...
          6070                6080                6090...

                    ... A A C A G A C T T T T T T A G A T A A T A C C A T C A A C C
                    ...               6100                6110                6120

                    tbpB
                    MET   LYS   HIS   ILE   ...
C A T C A G A G G A T T A T T T T A T G A A A C A C A T T C...
          6130                6140                6150...

                    ...PRO   LEU   THR   THR   LEU   CYS   VAL   ALA   ILE   SER   A
                    ... C T T T A A C C A C A C T G T G T G T G G C A A T C T C T G
                    ...               6160                6170                6180
```

FIG.8Z

```
LA  VAL  LEU  LEU  THR  ALA  CYS  GLY  GLY  SER  ...
C C G T C T T A T T A A C C G C T T G T G G T G G C A G T G...
            6190                6200               6210...
                    ...GLY  GLY  SER  ASN  PRO  PRO  ALA  PRO  THR  PRO  I
                    ... G T G G T T C A A A T C C A C C T G C T C C T A C G C C C A
                    ...             6220               6230               6240

LE  PRO  ASN  ALA  SER  GLY  SER  GLY  ASN  THR  ...
T T C C A A A T G C T A G C G G T T C A G G T A A T A C T G...
            6250                6260               6270...
                    ...GLY  ASN  THR  GLY  ASN  ALA  GLY  GLY  THR  ASP  A
                    ... G C A A C A C T G G T A A T G C T G G C G G T A C T G A T A
                    ...             6280               6290               6300

SN  THR  ALA  ASN  ALA  GLY  ASN  THR  GLY  GLY  ...
A T A C A G C C A A T G C A G G T A A T A C A G G C G G T A...
            6310                6320               6330...
                    ...THR  ASN  SER  GLY  THR  GLY  SER  ALA  ASN  THR  P
                    ... C A A A C T C T G G T A C A G G C A G T G C C A A C A C A C
                    ...             6340               6350               6360

RO  GLU  PRO  LYS  TYR  GLN  ASP  VAL  PRO  THR  ...
C A G A G C C A A A A T A T C A A G A T G T A C C A A C T G...
            6370                6380               6390...
                    ...GLU  LYS  ASN  GLU  LYS  ASP  LYS  VAL  SER  SER  I
                    ... A G A A A A A T G A A A A A G A T A A A G T T T C A T C C A
                          ...        6400               6410               6420
```

FIG.8A'

```
LE  GLN  GLU  PRO  ALA  MET  GLY  TYR  GLY  MET  ...
T T C A A G A A C C T G C C A T G G G T T A T G G C A T G G...
          6430                6440                6450...
                         ...ALA  LEU  SER  LYS  ILE  ASN  LEU  HIS  ASN  ARG  G
                         ... C T T T G A G T A A A A T T A A T C T A C A C A A C C G A C
                                         6460                6470                6480
                         ...

LN  ASP  THR  PRO  LEU  ASP  GLU  LYS  ASN  ILE  ...
A A G A C A C G C C A T T A G A T G A A A A A A A T A T C A...
          6490                6500                6510...
                         ...ILE  THR  LEU  ASP  GLY  LYS  LYS  GLN  VAL  ALA  G
                         ... T T A C C T T A G A C G G T A A A A A A C A A G T T G C A G
                                         6520                6530                6540
                         ...

LU  GLY  LYS  LYS  SER  PRO  LEU  PRO  PHE  SER  ...
A A G G T A A A A A A T C G C C A T T G C C A T T T T C G T...
          6550                6560                6570...
                         ...LEU  ASP  VAL  GLU  ASN  LYS  LEU  LEU  ASP  GLY  T
                         ... T A G A T G T A G A A A A T A A A T T G C T T G A T G G C T
                                         6580                6590                6600
                         ...

YR  ILE  ALA  LYS  MET  ASN  VAL  ALA  ASP  LYS  ...
A T A T A G C A A A A A T G A A T G T A G C G G A T A A A A...
          6610                6620                6630...
                         ...ASN  ALA  ILE  GLY  ASP  ARG  ILE  LYS  LYS  GLY  A
                         ... A T G C C A T T G G T G A C A G A A T T A A G A A A G G T A
                                         6640                6650                6660
                         ...
```

FIG.8B'

```
SN  LYS  GLU  ILE  SER  ASP  GLU  GLU  LEU  ALA  ...
A T A A A G A A A T C T C C G A T G A A G A A C T T G C C A...
              6670                6680                6690...
                                   ...LYS  GLN  ILE  LYS  GLU  ALA  VAL  ARG  LYS  SER  H
                                   ... A A C A A A T C A A A G A A G C T G T G C G T A A A A G C C
                                                  6700                6710                6720
                                   ...

IS  GLU  PHE  GLN  GLN  VAL  LEU  SER  SER  LEU  ...
A T G A G T T T C A G C A A G T A T T A T C A T C A C T G G...
              6730                6740                6750...
                                   ...GLU  ASN  LYS  ILE  PHE  HIS  SER  ASN  ASP  GLY  T
                                   ... A A A A C A A A A T T T T T C A T T C A A A T G A C G G A A
                                                  6760                6770                6780
                                   ...

HR  THR  LYS  ALA  THR  THR  ARG  ASP  LEU  LYS  ...
C A A C C A A A G C A A C C A C A C G A G A T T T A A A A T...
              6790                6800                6810...
                                   ...TYR  VAL  ASP  TYR  GLY  TYR  TYR  LEU  ALA  ASN  A
                                   ... A T G T T G A T T A T G G T T A C T A C T T G G C G A A T G
                                                  6820                6830                6840
                                   ...

SP  GLY  ASN  TYR  LEU  THR  VAL  LYS  THR  ASP  ...
A T G G C A A T T A T C T A A C C G T C A A A A C A G A C A...
              6850                6860                6870...
                                   ...LYS  LEU  TRP  ASN  LEU  GLY  PRO  VAL  GLY  GLY  V
                                   ... A A C T T T G G A A T T T A G G C C C T G T G G G T G G T G
                                                  6880                6890                6900
                                   ...
```

FIG.8C'

```
AL  PHE  TYR  ASN  GLY  THR  THR  THR  ALA  LYS   ...
T G T T T T A T A A T G G C A C A A C G A C C G C C A A A G...
            6910                6920                6930...
                                ...GLU  LEU  PRO  THR  GLN  ASP  ALA  VAL  LYS  TYR  L
                                ... A G T T G C C C A C A C A A G A T G C G G T C A A A T A T A
                                ...             6940                6950                6960

YS  GLY  HIS  TRP  ASP  PHE  MET  THR  ASP  VAL   ...
A A G G A C A T T G G G A C T T T A T G A C C G A T G T T G...
            6970                6980                6990...
                                ...ALA  ASN  ARG  ARG  ASN  ARG  PHE  SER  GLU  VAL  L
                                ... C C A A C A G A A G A A A C C G A T T T A G C G A A G T G A
                                ...             7000                7010                7020

YS  GLU  ASN  SER  GLN  ALA  GLY  TRP  TYR  TYR   ...
A A G A A A A C T C T C A A G C A G G C T G G T A T T A T G...
            7030                7040                7050...
                                ...GLY  ALA  SER  SER  LYS  ASP  GLU  TYR  ASN  ARG  L
                                ... G A G C A T C T T C A A A A G A T G A A T A C A A C C G C T
                                ...             7060                7070                7080

EU  LEU  THR  LYS  GLU  ASP  SER  ALA  PRO  ASP   ...
T A T T A A C T A A A G A A G A C T C T G C C C C T G A T G...
            7090                7100                7110...
                                ...GLY  HIS  SER  GLY  GLU  TYR  GLY  HIS  SER  SER  G
                                ... G T C A T A G C G G T G A A T A T G G C C A T A G C A G T G
                                ...             7120                7130                7140
```

FIG.8D'

```
LU  PHE THR VAL ASN PHE LYS GLU LYS LYS ...
AGTTTACTGTTAATTTTAAGGAAAAAAAAT...
      7150           7160         7170...
                     ...LEU THR GLY LYS LEU PHE SER ASN LEU GLN A
                     ... TAACAGGTAAGCTGTTTAGTAACCTACAAG
                     ...        7180       7190       7200

SP  ARG HIS LYS GLY ASN VAL THR LYS THR ...
ACCGCCATAAGGGCAATGTTACAAAAACCG...
      7210           7220         7230...
                     ...GLU ARG TYR ASP ILE ASP ALA ASN ILE HIS G
                     ... AACGCTATGACATCGATGCCAATATCCACG
                     ...        7240       7250       7260

LY  ASN ARG PHE ARG GLY SER ALA THR ALA ...
GCAACCGCTTCCGTGGCAGTGCCACCGCAA...
      7270           7280         7290...
                     ...SER ASN LYS ASN ASP THR SER LYS HIS PRO P
                     ... GCAATAAAAATGACACAAGCAAACACCCCT
                     ...        7300       7310       7320

HE  THR SER ASP ALA ASN ASN ARG LEU GLU ...
TTACCAGTGATGCCAACAATAGGCTAGAAG...
      7330           7340         7350...
                     ...GLY GLY PHE TYR GLY PRO LYS GLY GLU GLU L
                     ... GTGGTTTTTATGGGCCAAAAGGCGAGGAGC
                     ...        7360       7370       7380
```

FIG.8E'

```
EU   ALA  GLY  LYS  PHE  LEU  THR  ASN  ASP  ASN  ...
T G G C A G G T A A A T T C T T A A C C A A T G A C A A C A...
               7390           7400           7410...
                         ...LYS  LEU  PHE  GLY  VAL  PHE  GLY  ALA  LYS  ARG  G
                         ... A A C T C T T T G G C G T C T T T G G T G C T A A A C G A G
                         ...            7420           7430           7440

LU   SER  LYS  ALA  GLU  GLU  LYS  THR  GLU  ALA  ...
A G A G T A A A G C T G A G G A A A A A A C C G A A G C C A...
               7450           7460           7470...
                         ...ILE  LEU  ASP  ALA  TYR  ALA  LEU  GLY  THR  PHE  A
                         ... T C T T A G A T G C C T A T G C A C T T G G G A C A T T T A
                         ...            7480           7490           7500

SN   THR  SER  ASN  ALA  THR  THR  PHE  THR  PRO  ...
A T A C A A G T A A C G C A A C C A C A T T C A C C C C A T...
               7510           7520           7530...
                         ...PHE  THR  GLU  LYS  GLN  LEU  ASP  ASN  PHE  GLY  A
                         ... T T A C C G A A A A A C A A C T G G A T A A C T T T G G C A
                         ...            7540           7550           7560

SN   ALA  LYS  LYS  LEU  VAL  LEU  GLY  SER  THR  ...
A T G C C A A A A A A T T G G T C T T A G G T T C T A C C G...
               7570           7580           7590...
                         ...VAL  ILE  ASP  LEU  VAL  PRO  THR  ASP  ALA  THR  L
                         ... T C A T T G A T T T G G T G C C T A C T G A T G C C A C C A
                         ...            7600           7610           7620
```

FIG.8F'

```
YS   ASN   GLU   PHE   THR   LYS   ASP   LYS   PRO   GLU   ...
A A A A T G A A T T C A C C A A A G A C A A G C C A G A G T...
                  7630                 7640                 7650...
                                  ...SER   ALA   THR   ASN   GLU   ALA   GLY   GLU   THR   LEU   M
                                  ... C T G C C A C A A A C G A A G C G G G C G A G A C T T T G A
                                  ...                7660                 7670                 7680

ET   VAL   ASN   ASP   GLU   VAL   SER   VAL   LYS   THR   ...
T G G T G A A T G A T G A A G T T A G C G T C A A A A C C T...
                  7690                 7700                 7710...
                                  ...TYR   GLY   LYS   ASN   PHE   GLU   TYR   LEU   LYS   PHE   G
                                  ... A T G G C A A A A A C T T T G A A T A C C T A A A A T T T G
                                  ...                7720                 7730                 7740

LY   GLU   LEU   SER   ILE   GLY   GLY   SER   HIS   SER   ...
G T G A G C T T A G T A T C G G T G G T A G C C A T A G C G...
                  7750                 7760                 7770...
                                  ...VAL   PHE   LEU   GLN   GLY   GLU   ARG   THR   ALA   THR   T
                                  ... T C T T T T T A C A A G G C G A A C G C A C C G C T A C C A
                                  ...                7780                 7790                 7800

HR   GLY   GLU   LYS   ALA   VAL   PRO   THR   THR   GLY   ...
C A G G C G A G A A A G C C G T A C C A A C C A C A G G C A...
                  7810                 7820                 7830...
                                  ...THR   ALA   LYS   TYR   LEU   GLY   ASN   TRP   VAL   GLY   T
                                  ... C A G C C A A A T A T T T G G G G A A C T G G G T A G G A T
                                  ...                7840                 7850                 7860
```

FIG.8G'

```
YR  ILE  THR  GLY  LYS  ASP  THR  GLY  THR  GLY  ...
A C A T C A C A G G A A A G G A C A C A G G A A C G G G C A...
          7870                 7880                 7890...
                         ...THR  GLY  LYS  SER  PHE  THR  ASP  ALA  GLN  ASP  V
                         ... C A G G A A A A A G C T T T A C C G A T G C C C A A G A T G
                                       7900                 7910                 7920
                         ...

AL  ALA  ASP  PHE  ASP  ILE  ASP  PHE  GLY  ASN  ...
T T G C T G A T T T T G A C A T T G A T T T T G G A A A T A...
          7930                 7940                 7950...
                         ...LYS  SER  VAL  SER  GLY  LYS  LEU  ILE  THR  LYS  G
                         ... A A T C A G T C A G C G G T A A A C T T A T C A C C A A A G
                                       7960                 7970                 7980
                         ...

LY  ARG  GLN  ASP  PRO  VAL  PHE  SER  ILE  THR  ...
G C C G C C A A G A C C C T G T A T T T A G C A T C A C A G...
          7990                 8000                 8010...
                         ...GLY  GLN  ILE  ALA  GLY  ASN  GLY  TRP  THR  GLY  T
                         ... G T C A A A T C G C A G G C A A T G G C T G G A C A G G G A
                                       8020                 8030                 8040
                         ...

HR  ALA  SER  THR  THR  LYS  ALA  ASP  ALA  GLY  ...
C A G C C A G C A C C A C C A A A G C G G A C G C A G G A G...
          8050                 8060                 8070...
                         ...GLY  TYR  LYS  ILE  ASP  SER  SER  SER  THR  GLY  L
                         ... G C T A C A A G A T A G A T T C T A G C A G T A C A G G C A
                                       8080                 8090                 8100
                         ...
```

FIG.8H'

```
YS   SER  ILE  ALA  ILE  LYS  ASP  ALA  ASN  VAL  ...
A A T C C A T C G C C A T C A A A G A T G C C A A T G T T A...
          8110                8120                8130...
                              ...THR  GLY  GLY  PHE  TYR  GLY  PRO  ASN  ALA  ASN  G
                              ... C A G G G G G C T T T T A T G G T C C A A A T G C A A A C G
                              ...           8140                8150                8160

LU   MET  GLY  GLY  SER  PHE  THR  HIS  ASN  ALA  ...
A G A T G G G C G G G T C A T T T A C A C A C A A C G C C G...
          8170                8180                8190...
                              ...ASP  ASP  SER  LYS  ALA  SER  VAL  VAL  PHE  GLY  T
                              ... A T G A C A G C A A A G C C T C T G T G G T C T T T G G C A
                              ...           8200                8210                8220

HR   LYS  ARG  GLN  GLN  GLU  VAL  LYS  ***
C A A A A A G A C A A C A A G A A G T T A A G T A G T A A T...
          8230                8240                8250...

                              ... T T A A A C A C A A T G T T T G
                              ...           8260
```

# FIG.9A

Alignment of *M. catarrhalis* ORF3 proteins

```
         10        20        30        40        50
MLAFLIGAVMTITPVYTTFTPTKTPIKFFMAGLTFLIAHISHADDGRTDN
..................................................
                                                          60        70        80        90       100
                                                  ..P.......................G....T..................   Q8
                                                  QELINQEIATLEPIINHAQPELLSHDALTPKIEPILAQTPNPAEDTLIAD   4223

        110       120       130       140       150
EALLLDNPDLLNHALNSAVMTNHMAGVHALLPIYQKLPKDHQNGILLGYA
......................N...........................
                                                         160       170       180       190       200
                                                  NALAALDKGNAKKAIDELRRIIAIMPEYNVVRFHLARALFMDKQNEAALD   4223
                                                  ...V........A..G..................................   Q8

        210       220       230       240       250
QFNKLHADNLPEEVRQVVGQYRQALKQRDSWTWQVGMNLAKEDNINQTPK
...............R..................................
                                                         260       270       280       290       300
                                                  NTTQGQWTFDKPIDAITLSYQLGADKKWSLPKGAYVGANAQIYGKHHQNH   4223
                                                  ..................................................   Q8

        310       320       330       340       350
KKYNDHWGRLGANLGFADAKKDLSIETYGEKRFYGHERYTDTIGIRMSVD
................................................A.
                                                         360       370       380       390       400
                                                  YRINPKFQSLNAIDISRLTNHRTPRADSNNTLYSTSLIYYPNATRYYLLG   4223
                                                  ..................................................   Q8
```

FIG.9B

```
         410       420       430       440       450
ADFYDEKVPQDPSDSYQRRGIRTAWGQEWAGGLSSRAQISINKRHYQGAN
................E.................................
                   460       470       480       490       500
          LTSGGQIRHDKQMQASLSLWHRDIHKWGITPRLTISTNINKSNDIKANYH  4223
          ........Q.........................................  Q8

         510
KNQMFVEFSRIF*
............*
                                                                      4223
                                                                      Q8
```

# INTERN[redacted]NAL SEARCH REPORT

Int [redacted]pplication No
PCT/CA 99/00307

**A. CLASSIFICATION OF SUBJECT MATTER**
IPC 6 C07K14/79 C07K14/22 C12N15/31 C12Q1/68 A61K39/02
A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category ° | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | WO 97 32980 A (CONNAUGHT LAB ) 12 September 1997 (1997-09-12) page 3, line 29 -page 9, line 32 page 19, line 32 -page 30, line 6; examples 1-19 SEQ.ID.N.3 | 1-12 |
| A | WO 97 13785 A (CONNAUGHT LAB ;YANG YAN PING (CA); MYERS LISA E (CA); HARKNESS ROB) 17 April 1997 (1997-04-17) page 1, line 1 -page 3, line 8; examples 1-8 page 4, line 20 -page 8, line 31 | 1,9 |
| A | US 5 708 149 A (SCHRYVERS ANTHONY ET AL) 13 January 1998 (1998-01-13) abstract; figure 23 column 5, line 63 -column 6, line 28 | 1,6 |

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 15 October 1999 | 02/11/1999 |

| Name and mailing address of the ISA | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016 | Mateo Rosell, A.M. |

1

Form PCT/ISA/210 (second sheet) (July 1992)

International Application No
PCT/CA 99/00307

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ° | Citation of document, with indication,where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| A | SCHRYVERS A B ET AL: "COMPARATIVE ANALYSIS OF THE TRANSFERRIN AND LACTOFERRIN BINDING PROTEINS IN THE FAMILY NEISSERIACEAE"<br>CANADIAN JOURNAL OF MICROBIOLOGY,<br>vol. 35, no. 5, 1 May 1989 (1989-05-01),<br>pages 409-415, XP002020995<br>ISSN: 0008-4166<br>cited in the application<br>abstract<br>--- | 1 |
| A | RAONG-HUA YU ET AL: "THE INTERACTION BETWEEN HUMAN TRANSFERRIN AND TRANSFERRIN BINDING PROTEIN 2 FROM MORAXELLA (BRANHAMELLA) CATARRHALIS DIFFERS FROM THAT OF OTHER HUMAN PATHOGENS"<br>MICROBIAL PATHOGENESIS,<br>vol. 15, 1 January 1993 (1993-01-01),<br>pages 433-445, XP000612196<br>ISSN: 0882-4010<br>abstract<br>--- | 1 |
| P,X | MYERS L.E. ET AL., : "The transferrin binding protein B of moraxella catarrhalis elicits bactericidal antibodies and is a potential vaccine antigen"<br>INFECTION AND IMMUNITY,<br>vol. 66, no. 9, 1998, page 4183-4192<br>XP002118475<br>the whole document<br>----- | 2,7 |

1

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/CA 99/00307

**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.: 9
   because they relate to subject matter not required to be searched by this Authority, namely:
   Remark: Although claim 9 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
   because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
   because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International Application No
PCT/CA 99/00307

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| WO 9732980 A | 12-09-1997 | AU 1865397 A | 22-09-1997 |
| | | CA 2248095 A | 12-09-1997 |
| | | CN 1217748 A | 26-05-1999 |
| | | EP 0885300 A | 23-12-1998 |
| | | NZ 331777 A | 29-09-1999 |
| WO 9713785 A | 17-04-1997 | AU 7208296 A | 30-04-1997 |
| | | CA 2234409 A | 17-04-1997 |
| | | EP 0866803 A | 30-09-1998 |
| | | JP 11500744 T | 19-01-1999 |
| US 5708149 A | 13-01-1998 | US 5922562 A | 13-07-1999 |
| | | US 5922841 A | 13-07-1999 |
| | | US 5922323 A | 13-07-1999 |
| | | AU 705998 B | 03-06-1999 |
| | | AU 8102094 A | 29-05-1995 |
| | | BR 9408006 A | 03-12-1996 |
| | | CA 2175332 A | 18-05-1995 |
| | | WO 9513370 A | 18-05-1995 |
| | | CN 1141060 A | 22-01-1997 |
| | | EP 0728200 A | 28-08-1996 |
| | | JP 9506247 T | 24-06-1997 |
| | | NZ 275772 A | 27-04-1999 |